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

Olga Kellert, Daniele Panizza, Caterina Petrone

Olga Kellert, Daniele Panizza, Caterina Petrone

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

1. Abstract

This work investigates the role of prosody in the perception wh-exclamatives and (information seeking) wh-interrogatives in Cosenza Italian. We used reaction times (RTs) as a diagnostic of listeners' (in)security in sentence type disambiguation during a two-forced choice identification task. Our results show that listeners identify the two sentence types after the end of the utterance in most of the trials, and 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 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.

Keywords: Reaction Times, wh-exclamatives, wh-interrogatives, prosody, perception, Cosenza Italian

2. 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 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 for identifying exclamatives and interrogatives (cf. Batliner 1988; Eady & Cooper 1986; Sorianello 2011,2012; Gyuris & Mády & Szalontai 2013). However, it is still unclear to what extent temporally-distributed phonological/phonetic properties are exploited by listeners for the sentence type identification.

Consider the 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"). For instance, in the Italian variety spoken in Cosenza (Southern Italy), wh-exclamatives and wh-interrogatives contain different prosodic cues in both the prenuclear (i.e. at the beginning of the intonation contour) and in the nuclear (i.e., at the end of the intonation contour) regions. Wh-exclamatives exhibit a %H at the left edge of the intonational phrase, which is absent in wh-interrogatives (where a prenuclear H* is produced on the wh-constituent,Sorianello 2011 and section 3 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 wonder whether listeners rely on the nuclear information or whether the prenuclear region also contributes to the perception of sentence types. 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 about 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 that include pitch accents and edge tones (Pierrehumbert and Hirschberg 1990). Despite this composition 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 is aimed at investigating whether listeners rely solely on the information in the nuclear region or whether they can use prenuclear cues to identify the contrast between wh-exclamatives and wh-interrogatives.

Furthermore, we wonder whether the processing of the intonational contour is similar across sentence types or whether some cues are more salient in one specific sentence type. The %H is a marked pattern in Italian, since its use is restricted to few cases (like the wh-exclamatives). It also exhibits enhanced pitch excursion, which renders it perceptually salient (Sorianello, 2011). Hence, we hypothesize that the intonation of wh-exclamatives should be processed and identified much faster than the intonation of wh-interrogatives.

Concerning methodology, traditional identification tasks have been used in combination with reaction times in prosody research, especially within the categorical perception paradigm. (Chen 2003; Falé & Hub Faria 2006; Niebuhr 2007; Feldhausen & Pešková & Kireya & Gabriel 2011, among others). Reaction times are believed to reflect task difficulty (Massaro 1987). Reaction times have been used as a substitute of the discrimination scores to test whether the perception of prosodic contrasts is categorical or gradient (e.g., Chen 2003; Niebuhr 2007). This methodology has been applied to contrasts at the level of pitch accents (e.g., Chen 2003; Niebuhr 2007; Feldhausen & Pešková & Kireya & Gabriel 2011), prosodic boundaries (e.g., Schneider 2011; Petrone & Truckenbrodt & Wellmann & Holzgrefe-Lang & Wartenburger & Höhle 2017) and global phonetic cues (such as pitch range, see Borràs-Comes & Vanrell & Prieto 2010). These studies are based on manipulated stimuli. Given a continuum of manipulated stimuli, it is generally supposed that, in case of categorical boundary perception, reaction times should be longer at the location of the continuum corresponding to the category boundary, while they should be shorter at the other locations of the continuum. On the other hand, a gradual increase of RTs might reflect a gradual increase of the ambiguity of the stimuli and would point to a more gradient boundary perception.

In this study, we also employ an experimental design that involves an identification task with measurement of reaction times. However, unlike previous studies, we will measure identification

scores and reaction times in response to natural (and not manipulated) stimuli as a first step into the investigation of this understudied contrast.

The structure of the paper is as follows: after a short review of the literature on wh-exclamatives and wh-interrogatives (Section 3), the goals of the study (Section 4) and the description of the experiment (Sections 5 and 6) will be reported. Finally, discussion of the findings and future perspectives will be illustrated (Section 7).

3. Previous investigations of the prosody of wh-exclamatives and wh-interrogatives

Since Italian is the language under investigation here, we will mainly focus on the prosody of Italian wh-exclamatives and wh-interrogatives (but see, e.g., Vanrell 2007; Hedberg, Sosa, Görgülü & Mameni 2010 for studies onwh interrogatives in other languages). In Italian, one of the most typical intonational characteristics of wh-interrogatives is the presence of utterance-final f0 fall (e.g., Chapallaz 1964; Avesani 1995; Sorianello 2011). However, the intonational pattern of wh-interrogatives might also differ in many respects depending either on the pragmatic function of the wh-interrogative (e.g. rhetorical function vs. information seeking function, see Hedberg & Sosa & Görgülü & Mameni 2010) or on the regional variety (see Avesani 1995; Sorianello 2011; Gili-Fivela 2015, among others). In particular, in Cosenza Italian (i.e., the Italian variety under investigation here), the nuclear accent of information-seeking wh-interrogatives is associated with a rising high tone L+H* while the falling contour at the end of the utterance is represented by a low boundary tone L% (see Sorianello 2011; Sorianello,Giordano & Petrone 2011).

The prosody of wh-exclamatives has not been studied very extensively so far (see Batliner 1988; Sorianello 2011; Gyuris & Médy & Szalontai 2013). Studies so far have mostly focused on the syntax and semantics of the difference between wh-exclamatives and wh-interrogatives (see Portner & Zanuttini 2003; Castroviejo 2006, among others). According to these studies, wh-exclamatives and wh-interrogatives differ with respect to factivity, i.e., only exclamatives but not interrogatives imply a true proposition. Portner and Zanuttini (2003) derive the factivity of wh-exclamatives from two factors. Exclamatives (including wh-exclamatives) can only be embedded under factive predicates (Portner & Zanuttini 2003: 46):

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

Wh-exclamatives can never be used as questions, and they 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 both wh- and non-wh-exclamatives from other sentence types. Cross-linguistically, exclamatives have been described as being characterized by an initial extra high pitch followed by a falling intonation contour (see O'Connor and Arnold 1961 for English; Delattre 1966 for French; D'Eugenio 1976 for Italian; Batliner 1988 for German). Other languages such as Hungarian are rather characterized by an initial low pitch (a low boundary tone) (see Gyuris & Mády & Szalontai 2013). Concerning the nuclear region, Gyuris & Mády & Szalontai (2013) found out that Hungarian listeners classified acoustic stimuli as exclamatives and not as interrogatives if the stimuli 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. The two sentence types are characterized by the same ending, i.e., a L-L% tonal sequence (see figure 1 and 2).[1] However, they are differentiated at a phonological level in the regions preceding the utterance's final falls, 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 the presence of 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 & D'Imperio & Savino & Avesani 2005). Wh-interrogatives are characterized instead by a prenuclear H* on the wh-constituent. Moreover, the nuclear accent is specified differently, i.e., a low tone L* characterizes wh-exclamatives and a low-high tone L+H* appears in wh-interrogatives. Along with the intonational cues, wh-exclamatives are also characterized by an extra lengthening of the nuclear stressed syllable duration (Sorianello 2011). These differences are illustrated in Figures 1 and 2.

While there are still no perception studies on wh-exclamatives in Cosenza Italian, previous studies have suggested that the prenuclear region can contribute to the identification of non-wh-exclamatives in other Southern varieties of Italian. 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 investigating 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 application to intonational 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. It should be noted, however, that these kinds of studies (e.g., using the gating task) are difficult to implement, and they do not allow tracking of the timing of utterance interpretation.

Insert figure 1 here

Figure 1: Quanti pesci hai preso! "How many fishes you took!"(Sorianello 2011:316)

Insert figure 2 here

Figure 2: Information seeking wh-interrogative Che cosa le regalerebbero? "What would they give her as a present?" (Gili Fivela 2015:182)

4. 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 contrast in both sentence types as described in section 3, listeners should be capable of distinguishing wh-interrogatives from wh-exclamatives on the basis of prosody (Hypothesis 1).

Our second goal is 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 noticable already in the prenuclear region (Sorianello 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 seek to investigate any 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

[1] Other types of wh-interrogatives that do not start with quanti,e 'how many' (e.g. Chi le vendeva? 'Who was selling them?') may show a high boundary tone H% at the end of the utterance (see Sorianello & Giordano & Petrone 2011).

high boundary tone right at the beginning of the utterance (i.e. %H at the left edge of the intonational phrase) (see Sorianello 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 an 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 two-alternatives forced-choice identification task.

5. Experiment

5.1. Experimental stimuli

We constructed 20 morpho-syntactically and lexically identical wh-exclamatives and wh-interrogatives (see e.g., (1)). Each target sentence contained the following syntactic structure: a complex wh-constituent (i.e., *quanti* 'how many' + noun); a verb phrase consisting of an auxiliary and a past participle; 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 of 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 a pragmatic context that only matched one or the other sentence type. Below we list examples of a wh-exclamative context and a wh-interrogative context:

(6) wh-exclamative context: your mother told you that her friend has spent 10 years of her life to write novels and shows to you the 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 told you that her friend has spent 10 years of her life to write 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.[2] No instructions were given as to what specific prosodic pattern to use in sentence production. The 20 target sentences were presented randomly and interspersed among 20 fillers of

[2] The discourse context was only presented in the production of the stimuli to elicit a specific intonation contour not in the identification task because otherwise the discourse context would bias the responses of the listeners.

non-wh-interrogatives and non-wh-exclamatives like yes-no interrogatives, imperatives and declaratives (e.g., yes-no interrogative: "Do you like coffee?" Imperative: "Open the door!" declarative: "I came late today"). A complete list of experimental target sentences and the fillers is given in the Appendix.

For the perception experiment, wh-exclamatives and wh-interrogatives constituted our auditory target stimuli. For each sentence, we marked 4 critical points or "marks" at sequential temporal locations (M1-M4): M1= the beginning of the utterance; M2= end of the wh-constituent; M3= end of the verb phrase and M4= end of the subject constituent which was also the end of the utterance (M2-M4 are represented by closing brackets "]" and M1 by an opening bracket "[“):

(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, included the wh-constituent (e.g., *Quanti romanzi*); the middle region, between M2 and M3, included the verb phrase (e.g., *ha scritto*); and the final region, from M3 to M4, included the subject phrase at the end of the utterance (e.g., *la tua amica*).

5.2. Tasks and procedure

Since it has been shown that there are durational differences across the two sentence types (cf. Sorianello 2011), which was also clear from an initial analysis of our data, we measured the duration of each region. Statistical analysis of the duration of the three regions showed that the initial region was significantly longer in the exclamative condition than in the interrogative condition (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 (inter. = 751 ms vs. excl.= 933ms, $p<.001$). Given that the duration of each region seemed to play a crucial role in identification, this parameter was included as a variable in the statistical models (see section 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 Università della Calabria. They reported no hearing problems.

Listeners that participated in the experiment had 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 the key as soon as they were certain of the sentence type.

This instruction was made 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 middle region). Before the experiment, listeners had a short practice session, with four practice trials for sentence type identification. Listeners 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).[3]

Each trial began with a written question to the participants, asking if they were ready to start. A

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

beep was used to signal that an utterance was about to start in order to draw participant's attention to the stimulus. The session began with the presentation of the auditory stimulus. For each listener, the identification task included 2 blocks (i.e., the stimuli were divided into two parts). The goal of this manipulation was to check whether repetition of the same sentences influenced the reactions times obtained from the identification task (Bentin & McCarthy 1994). Block 1 contained 10 wh-exclamative sentences and 10 wh-interrogative sentences that were not lexically identical ("non-minimal pair condition", e.g., "How many poems your professor has written!" vs. "How many cigarettes does our father smoke?"). Stimuli were presented in random order and interspersed among 10 fillers of non-wh-exclamatives and non-wh-interrogatives. The first block of the experiment thus contained 30 sentences in total. Block 2 also contained 10 wh-exclamatives, 10 wh-interrogatives and 10 fillers. This time, 5 of the 10 wh-exclamatives and 5 of the 10 wh-interrogatives were lexically identical to the 5 wh-interrogatives and 5 wh-exclamatives presented in the first block. In other words, each participant heard the same sentence under both the exclamative and the interrogative condition (i.e. "minimal pair" condition, e.g., "How many novels has your friend written? " vs. "How many novels your friend has written! "). In total, each participant heard 60 sentences (i.e., 20 exclamatives, 20 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.

6. Statistical analysis and results

Before going into detail regarding the statistical analysis, we summarize our results with respect to the goals and the hypotheses we formulated in section 4. 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 both sentence types before they hear the end of the sentence but this pattern is very rare in our data; they distinguish both sentence types much more often 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 across sentence types, so that the processing advantage for wh-exclamatives disappears when the segmental duration is taken into account.

6.1. Identification task

The 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 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, suggestingthat some listeners are 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 judging wh-exclamatives than in judging wh-interrogatives. However, the difference between correct and incorrect early responses could not be assessed statistically because of the low

number of observations.

We run a generalized linear mixed model (GLMM) in order to analyze the accuracy of late responses, where we adopted *sentence type* (interrogatives vs. exclamative) and *block* (block 1 vs. 2) as fixed factors and *item* and *participant* as random factors with maximal random effect structure (cf. Barr & Levy & Scheepers & Tily 2013), that is, with the greatest 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 ($\beta = .11$, $z = .32$, $p=.75$).The factor *block* was also not significant ($\beta = .26$, $z= .81$, $p=.42$), 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* (interrogatives vs. exclamative) and *response time* (early vs. late responses) as fixed factors. The model confirmed no significant effects for *sentence type* for early responses (estimate = 1.65, $z = 1.24$, $p=.21$).

	Condition exclamatives		Condition interrogatives	
	Correct responses	Incorrect responses	Correct responses	Incorrect responses
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

Table 1: responses given in each time region.

6.2. Reaction Times (RTs)

We ran a statistical analysis on the reaction times 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 reaction times to achieve a normal distribution (cf. Baayen 2008). The dependent variable was the reaction time 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 achieved by applying a Linear Mixed Model (LMM) to the reaction times where we adopted *sentence type* as the main factor of interest and *item* and *participant* as random factors with maximal random effects structure (cf. Barr & Levy & Scheepers & Tily 2013). We checked whether the factor *block* had any effect on reaction times. A LMM model showed a non-significant effect of *block*, i.e., repetition of the sentences did not have any influence on the reaction times ($\beta = -0.01$, $t=-0.16$, $p=.87$). Furthermore, there was no interaction between *sentence type* and *block* ($\beta = 0.03$, $t=-.016$, $p=.48$). 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 wh-interrogatives. In the exclamative condition, listeners took on average 525 ms 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 (β = 0.07, t=3.49, p=<.01). In the next round of analyses, we checked whether reaction times 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 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. To investigate whether these differences were statistically significant we conducted a LMM adopting *sentence type* (wh-exclamative vs. wh-interrogative) and *response type* (early vs. late responses) as fixed factors and listeners and items as random factors, with random slopes and intercepts. The effect of *response type* was significant as expected (β=0.06, t=3.52, p=<.01). The effect of *sentence type* was also significant for late responses (β=1.2, t=19.42, p=<.01). This means that RTs differed significantly with different sentence types and different response types (early vs. late responses). However, the interaction between *sentence type* and *response type* (i.e. early responses) was not significant (β=0.11, t=0.83, p=.42). This last result should be considered with caution in that the number of observations regarding early responses was very low (see Table 1). Figure 3 presents reaction times across sentence type, in separate plots for early and late responses. As it can be seen, reaction times are different for late responses, while they are very similar for early responses:

Insert figure 3 here

Figure 3: log-transformed reaction times across sentence type and response type. The "0" value in the y-axis represent the end of the utterance, positive values represent RTs after the end of the utterance and negative RTs represent values before the end of the utterance.

Given that the duration of the stimuli was not balanced across the two sentence types (see Section 5.1), we conducted another analysis that included the duration of the initial region (containing the wh-phrase), the duration of the middle region (containing the verb phrase) and that of the final region (containing the subject phrase) as covariates. This analysis addressed the question of whether the difference in reaction times between wh-exclamatives and wh-interrogatives revealed by the previous analyses was caused by *sentence type* or if it was an epiphenomenon resulting from the durational difference of the segmental regions of the wh-exclamatives vs. wh-interrogatives. The LMM including *sentence type, response time,* duration of the *initial region, middle region* and that of the *final region* as fixed factors and with *item* and *factor* as random factors yields the following results. There was no significant difference across *sentence type* (β = 0.01, t=-0.77, p=.44) nor was there any interaction between *sentence type* and *response type* (β = 0.1, t=-1.08, p=.28). Instead, *response type* (β = 1.2, t=21.9, p<.001), *initial region* (β = -0.31, t=-4.15, p<.001), *middle region* (β = -0.23, t=-2.56, p<.02) and *final region* (β = -0.33, t=-5.51, p<.001) were 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 response times while the main factor of our experimental design, i.e. the *sentence type*, is not significant.

7. 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 in more than 90% of trials for both wh-exclamatives and wh-interrogatives indicates that there must be some prosodic marker which guided listeners' judgments. This

experiment is 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 the pragmatic interpretation. The fact that our experiment elicited early responses roughly to the same degree in wh-exclamatives (8.3%) and wh-interrogatives (9%) is compatible with the hypothesis that listeners could either employ prosodic information contained either in the nuclear 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) only one listener in one trial was able to identify a sentence in the region immediately following nuclear cues (i.e., the middle region); c) the early responses were provided on average 200 ms before the end of the utterance and more than 2 seconds after the offset of the region containing the nuclear cues.; and d) the final region significantly affect the RTs.

However, the fact that the duration of the initial and middle regions also significantly affected the RTs strongly might be indicative that prosodic information in the prenuclear section were also exploited by the listeners for identifying the 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 our instructions to be both fast and accurate, listeners might have accumulated different phonetic/phonological cues while listening to the utterance in order to augment the probability for a reliable response.[4] The fact that listeners gave responses after the end of the utterance does not rule out the possibility that some phonetic/phonological cues for sentence type disambiguation are located at the beginning of the utterance in the prenuclear position. We have chosen the identification task combined with RT measurement because it is a simple technique which was well suited for a preliminary investigation of perceptual differences in wh-exclamatives and wh-interrogatives. However, further improvement of the methodology as well as stimuli selection is needed.

Concerning the stimuli selection, the current experiment is based on natural stimuli and does not allow us to distinguish which acoustic marker contributed to the relatively well performed judgments on sentence type disambiguation at the end of the utterance. The speaker produced wh-exclamatives and wh-interrogatives based on specific pragmatic contexts, yet she was not asked to produce a specific set of intonation contours.

In future studies, we will investigate the influence of the 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).

Another potential problem with measuring RTs in an identification task where listeners have to press a button is that it contains a time delay between sentence type identification and the response reaction (i.e., pressing the button in our experiment). To address this issue, alternative methods for registering the response reaction could be implemented. For instance, it has been proposed that visual recognition might diminish the response delay significantly and thus might be an alternative task for sentence type recognition (Carreiras & Armstrong & Perea & Frost 2014). Additionally, we could adopt on-line measures like eye movements or mouse tracking (Marslen-Wilson & Tyler & Warren & Grenier & Lee 1992, Pynte & Prieur 1996, Tomlinson & Fox Tree 2011, Warren 2014).

To conclude, our study have shown first important results of concerning the identification and processing of wh-exclamatives and wh-interrogatives on the basis of prosody in Cosenza Italian. However, listeners wait until the end of the utterance to respond in most of the trials. Furthermore, our study shows that wh-exclamatives are processed faster than wh-interrogatives, but this effect

[4] We thank a reviewer for pointing out this hypothesis to us.

disappears when the duration of different regions of the utterance is taken into account.

8. Appendix_Data Set

Item List_Target sentences	M1[5]	M2	M3	M4
Item List_Target sentences of Interrogatives	Placement of markers in interrogatives on the durational level (measured in ms)			
Quanti romanzi ha scritto la tua amica	1.052	1.733	2.252	3.127
Quanti libri ha pubblicato il tuo professore	1.367	1.828	2.507	3.596
Quante sigarette ha fumato papa	1.219	1.926	2.595	3.074
Quanti paesi ha visto tua sorella	1.141	1.651	2.138	2.847
Quante cose ha aggiustato tuo padre	1.421	1.996	2.587	3.169
Quante birre ha bevuto la tua amica	0.999	1.490	2.052	2.747
Quanti chili ha perso tuo nipote	0.895	1.351	1.870	2.613
Quanti corsi ha seguito tua sorella	1.094	1.577	2.222	3.069
Quanta torta ha mangiato tua sorella	1.090	1.618	2.222	3.059
Quanti libri ha comprato tuo padre	0.809	1.285	1.965	2.710
Quanti soldi ti ha dato tuo padre	1.186	1.778	2.280	2.935
Quanti vestiti ha disegnato il tuo amico	1.049	1.684	2.285	3.100
Quanti pesci ha pescato tuo fratello	1.130	1.626	2.249	3.113
Quanti cd ha inciso tuo zio	1.124	1.739	2.256	3.013
Quante arance ha raccolto tuo nonno	1.335	1.860	2.494	3.169
Quanti quadri ha dipinto tua zia	1.197	1.688	2.325	3.055
Quanti dolci ha preparato tua madre	1.230	1.751	2.402	3.149
Quanti fiori ha piantato tua nonna	1.413	1.887	2.492	3.255
Quante farfalle ha catturato tuo fratello	0.624	1.266	1.926	2.759
Quante scarpe ha comprato tua zia	1.219	1.780	2.476	3.063
Item List_Target sentences of exclamatives	Placement of markers exclamatives on the durational level (measured in ms)			
Quanti romanzi ha scritto la tua amica	1.286	1.985	2.575	3.438
Quanti libri ha pubblicato il tuo professore	1.087	1.655	2.356	3.570
Quante sigarette ha fumato papà	1.098	1.960	2.737	3.348
Quanti paesi ha visto tua sorella	1.076	1.600	2.062	2.970
Quante cose ha aggiustato tuo padre	1.030	1.553	2.235	3.044
Quante birre ha bevuto la tua amica	1.097	1.619	2.174	3.180
Quanti chili ha perso tuo nipote	0.891	1.307	1.800	2.673

[5] The numbers under M1 correspond to measurements of the duration from the beginning of the acoustic file to the beginning of the utterance.

Quanti corsi ha seguito tua sorella	1.207	1.766	2.237	3.237
Quanta torta ha mangiato tua sorella	1.068	1.645	2.158	3.032
Quanti libri ha comprato tuo padre	0.686	1.420	2.068	3.147
Quanti soldi ti ha dato tuo padre	1.037	1.803	2.265	3.062
Quanti vestiti ha disegnato il tuo amico	0.942	1.828	2.454	3.361
Quanti pesci ha pescato tuo fratello	2.031	2.638	3.213	4.245
Quanti cd ha inciso tuo zio	1.143	1.886	2.346	3.226
Quante arance ha raccolto tuo nonno	1.750	2.301	2.884	3.789
Quanti quadri ha dipinto tua zia	1.711	2.214	2.828	3.753
Quanti dolci ha preparato tua madre	1.482	2.028	2.671	3.680
Quanti fiori ha piantato tua nonna	1.335	1.802	2.463	3.486
Quante farfalle ha catturato tuo fratello	0.803	1.457	2.129	3.141
Quante scarpe ha comprato tua zia	1.054	1.670	2.316	3.240

Itemlist_Fillers

Vieni stasera?
Mi fai un caffè?
Apriresti la finestra?
Piove tanto?
bella?
Mi daresti il tuo numero?
Perch_ piangi?
Hai visto il mio ragazzo?
Hai 25 anni?
Sei una stronza!
C‘è qualcuno al telefono!
Forse hai ragione!
Sei una persona speciale!
Vieni sta sera!
Guarda 'sto video!
Sei bellissima!

9. References

Avesani, Cinzia. 1995. ToBIt. 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.), XXII, Povo (TN), Italy, 85–98.

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.

Baayen, Harald. 2008. *Analyzing linguistic data. A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Borràs-Comes, Joan & Maria del Mar Vanrell & Pilar Prieto. 2010. The role of pitch range in establishing intonational contrasts in Catalan. *Fifth International Conference on Speech Prosody,* Chicago 11–14 May 2010. Paper 100103. 1–4.

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*, vol. 20(1), 130–149.

Castroviejo, Elena. 2006. *Wh-Exclamatives in Catalan*. Universitat de Barcelona. (Doctoral dissertation.)

Carreiras, Manuel & Blair C. Armstrong & Manuel Perea & Ram Frost. 2014. The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences (TICS)* 18(2), 90–98.

Chapallaz, Marguerite. 1964. Notes on the Intonation of Questions in Italian. In David. Abercrombie & D.B. Fry & P.A.D. MacCarth &, 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*, Geneva, 97–100.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Delattre, P .1966. Les 10 intonations de base du français, *The French Review*, vol. 40(1), 1–14.

D'Eugenio, Antonio. 1976. The intonation systems of Italian and English, *Rassegna Italiana di Linguistica Applicata* 8(1), 57–85.

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 Speech Prosody 2006*, 3rd International Conference Dresden, 69–72.

Feldhausen, Ingo & Andrea Pešková & Elena Kireva & Christoph Gabriel. 2011. Categorical perception of Porteño nuclear accents. *Proceedings of the 17th International Congress of Phonetic Sciences 2011*, Hong Kong, 116–119.

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 Phrasin,*. 362–389. Oxford: Oxford University Press.

Gili-Fivela, Barbara & Cinzia Avesani & Marco Barone & Giuliano Bocci & Claudia Crocco & Mariapaola D'Imperio & Rosa Giordano & Giovanna Marottaa & Michelina Savino& Patrizia Sorianello. 2015. Intonational phonology of the regional varieties of Italian. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 140–197. Oxford: Oxford University Press.

Grosjean, François. 1980. Spoken word recognition processes and the gating paradigm, *Perception & Psychophysics*, vol. 28, 267–283.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2006. Experimental approaches to establishing discreteness of intonational contrasts. In Stefan Suhoff & Denisa Lenertovà & Roland Meyer & Sandra Pappert & Petra

Augurzky & Ina Mleinek & Nicole Richter & Johannes Schließer (eds.), *Methods in Empirical Prosody Research*, 321–334. Berlin: Mouton de Gruyter.

Gyuris, Beáta & Katalin Mády & Ádám Szalontai. 2013. *Experimental investigations on the prosody of Hungarian exclamatives*. (Ms.)

Halliday, Michael A. K. 1985. *An introduction to Functional Grammar*, London: Edward Arnold.

Hedberg, Nancy& Juan M. Sosa & Emrah Görgülü & Morgan Mameni. 2010. The prosody and meaning of wh-questions in American English. In *Proceedings of Speech Prosody* 2010. Chicago, Illinois.

Hyman, Ray. 1953. Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology* 45, 188–196.

König, Ekkehard & Peter Siemund. 2007. Speech act distinctions in grammar. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. I. 2nd edition, 276–324. Cambridge: Cambridge University Press.

Ladd, D. Robert & Rachel Morton. 1997. The perception of intonational emphasis. Continuous or categorical? *Journal of Phonetics* 25, 313–342.

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 D. & 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*, London: Erlbaum.

Niebuhr, Oliver. 2007. Categorical perception in intonation. A matter of signal dynamics? *Proceedings of the Interspeech 2007*, Antwerp, Belgium, 109-112.

O'Connor, Joseph D. & Gordon F. Arnold. 1961. Intonation of Colloquial English, London: Longman.

Peirce, Jonathan W. 2007. PsychoPy – Psychophysics software in Python, *Journal of Neuroscience Methods* 162(1-2), 8–13.

Petrone, Caterina. 2008. *From targets to tunes. Nuclear and prenuclear contribution in the identification of intonation contours in Italian*. Laboratoire de Parole et Langage, Université de Provence (Doctoral dissertation.)

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*, 207–230. Berlin: Springer Verlag.

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 & 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.

Pisoni, David B. & Jeffrey Tash. 1974. Reaction times to comparisons within and across phonetic categories, *Perception & Psychophysics*, vol. 15(2), 285–290.

Portner, Paul & Rafaela Zanuttini. 2003. Exclamative clauses. At the syntaxsemantics interface. *Language* 79(1), 39–81.

Pynte, Joel & Benedicte Prieur. 1996. Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes*, 11, 165–191.

Savino, Michelina & Martine Grice. 2008. *Reaction time in the perception of intonational contrasts in Italian*. Paper presented at the Third TIE Conference on Tone and Intonation, Lisbon, 15-17 September 2008.

Sadock, Jerrold M. & Arnold M. Zwicky. 1985. Speech acts distinctions in syntax. In Timothy Shopen (ed.), *Language typology and syntactic description*, 155–196. Cambridge: Cambridge University Press.

Schneider, Edgar W. 2011. *English around the world. An introduction*. Cambridge: Cambridge University Press.

Solé, Maria-Josep & Patrice Speeter Beddor & Manjari Ohala. 2007. *Experimental approaches to phonology*. Oxford: Oxford University Press.

Sorianello, Patrizia. 2001. Modelli intonativi dell'interrogazione in una varietà di italiano meridionale. *Rivista Italiana di Dialettologia*, vol. 25, 85–108.

Sorianello, Patrizia. 2011. Aspetti prosodici e pragmatici dell'atto esclamativo. *Studi Linguistici e Filologici Online* 9, 287–332.

Sorianello, Patrizia & Riccardo Giordano & Caterina Petrone. 2011. *L'Italiano parlato a Cosenza*. PPt presented at the workshop ToBI Italiano CRIL, Lecce 29/01/2011.

Sorianello, Patrizia. 2012. A prosodic account of Italian exclamative sentences. A gating test. In Qiuwu & Hongwei Ding & Daniel Hirst (eds.), *The Proceedings of the 6th International Conference on Speech Prosody*, Shanghai (China), May 22-25, vol. I, 298–301. Tongji University Press,

Suhoff, Stefan & Denisa Lenertovà & Roland Meyer & Sandra Pappert & Petra Augurzky & Ina Mleinek & Nicole Richter & Johannes Schließer (eds.), *Methods in Empirical Prosody Research*. Berlin: Mouton de Gruyter.

Tomlinson, John M. & Jean E. Fox Tree. 2011. Listeners' comprehension of uptalk in spontaneous speech. *Cognition* 119(1), 58–69.

Vanrell, Maria del Mar. 2007. A tonal scaling contrast in Majorcan Catalan interrogatives. In Gorka Elordieta & Marina Vigário (eds.), *Journal of Portuguese Linguistics* (special issue on Prosody of Iberian Languages), vol. 6.1, 147–178.

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: ASSTA.

Figure 1

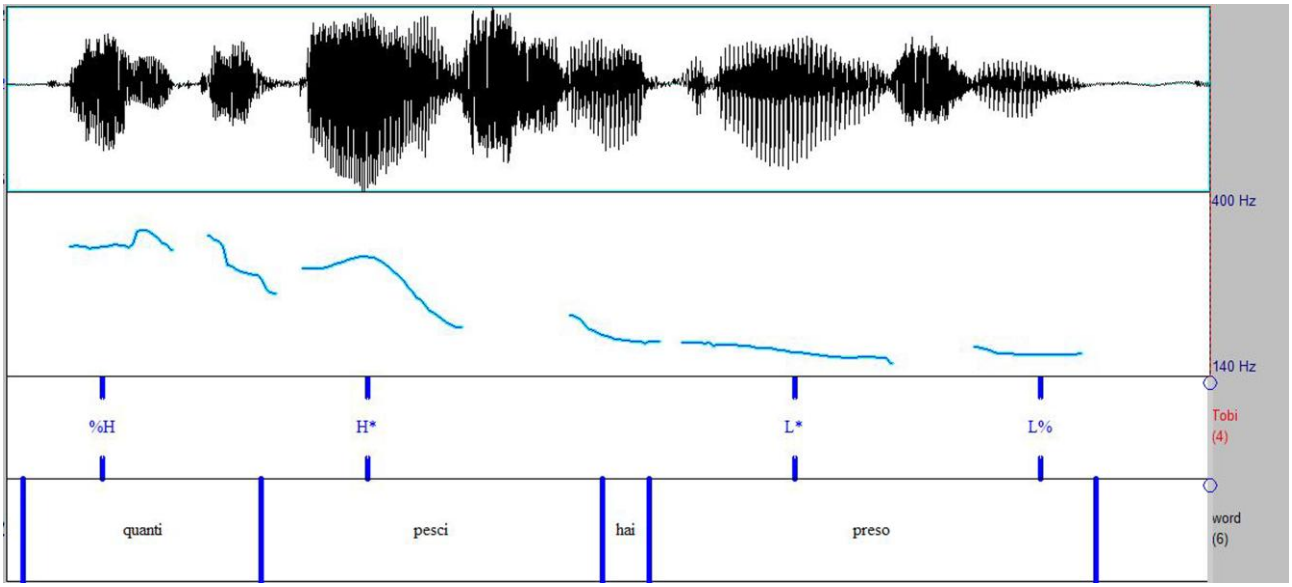


Figure 2

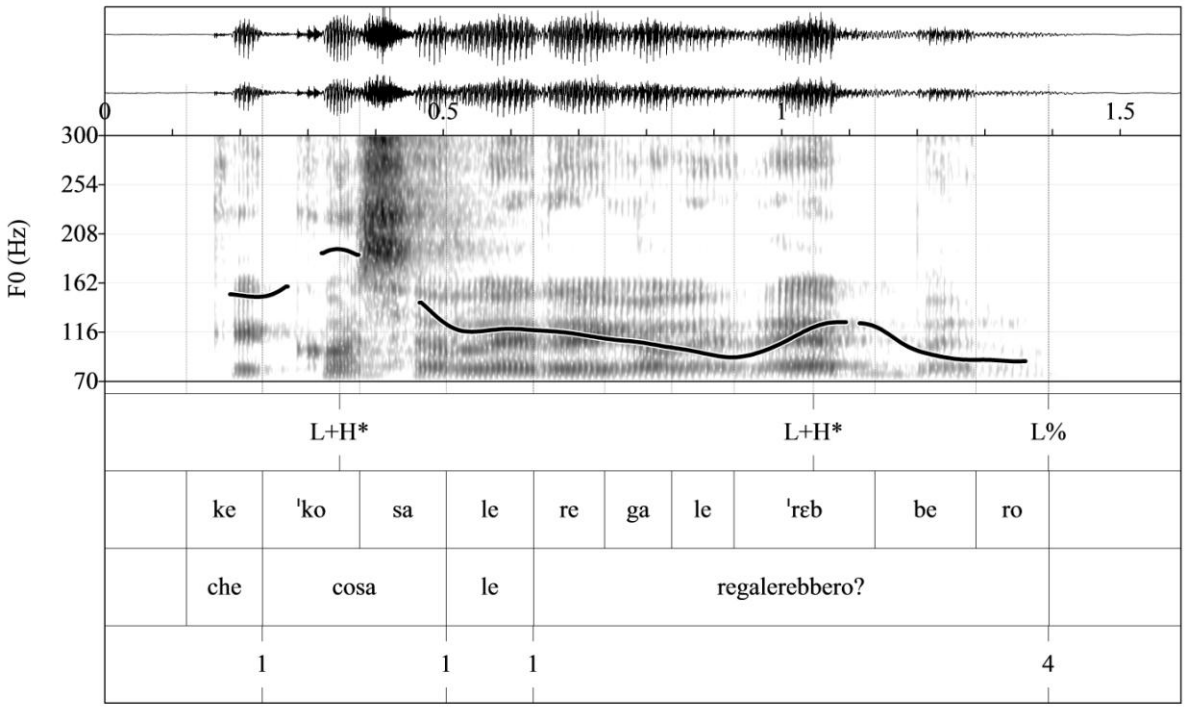
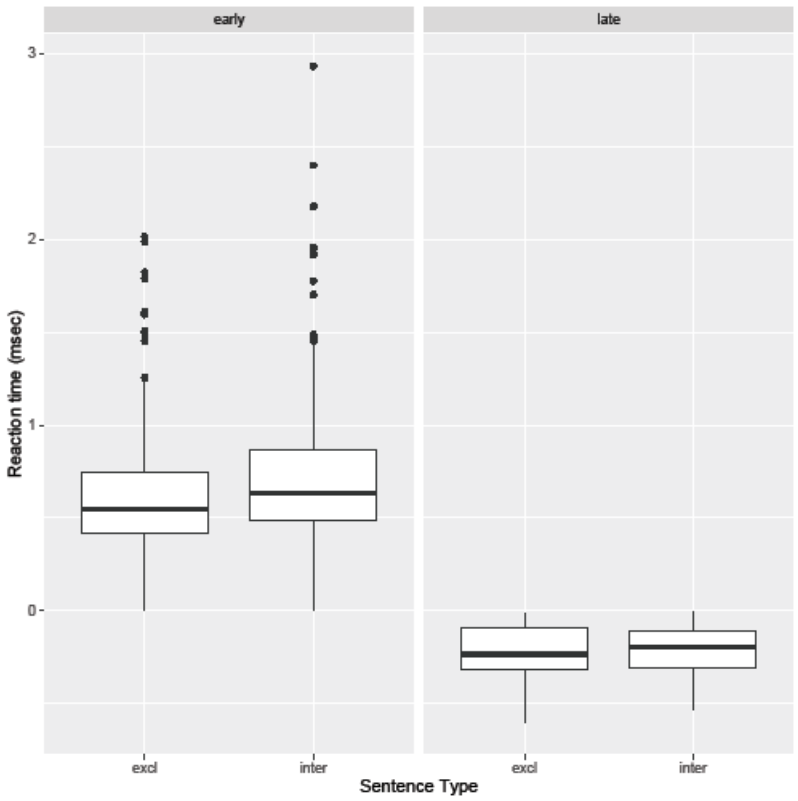


Figure 3


early
late
3
2
1
0
Reaction time (msec)
excl
inter
excl
inter
Sentence Type