Reaction Time Experimentation in Prosody Research

Abstract

This paper reports a perceptual study using an identification task in which we investigated the prosody of two different sentence types (namely exclamatives and interrogatives) which happen to be homophone in Italian Quanti romanzi ha scritto la tua amica 'How many novels your friend wrote!' or 'How many novels did your friend write?'. Unlike previous experiments on sentence type disambiguation (see Sorianello 2011, Gyuris et al 2013), the present study employs an on-line method using the *R*eaction *T*ime (RT) measurement, in addition to the measurement of response frequencies. Using this method may help us to reveal the location of prosodic features that trigger listener's judgements on sentence type disambiguation and to characterize their (in)security in responses to acoustic stimuli. Our results show that subjects identify both sentence types after the end of the utterance and not before it and that the mean RTs are equally long under both sentence type conditions. We conclude from these results that the prosodic cues that do matter for perception are linked to the final region of the utterance in both sentence type conditions. Given that subjects take the same amount of time to identify both sentence types, we conclude that participants of the experiment are equally secure about their responses under both conditions and find the identification task equally difficult with both sentence types.

Keywords: Reaction Time measurement, sentence type identification, exclamatives, interrogatives, prosody, perception, Italian

1. Introduction

Many languages present sentences that are ambiguous between an exclamative and an interrogative interpretation. In Italian, for instance, wh-interrogatives and wh-exclamatives including the wh-constituent quanto ('how much' in English) or quanti ('how many') are morpho-syntactically identical and they are ambiguous between the two interpretations outlined in (1) :

(1) Italian

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 the literature, there are several studies that have demonstrated that listeners rely on the prosodic information to disambiguate between sentence types like declaratives and yes/no questions (cf. Batliner 1988, Eady & Cooper 1986, Sorianello 2012, among many others).

For instance, Sorianello (2012) employed a gating experiment in which stimuli of partial or complete utterances in Bari Italian were presented to listeners in order to verify the melodic differences between non-wh-exclamatives vs. assertives (e.g. Piove! 'It is raining!' vs. Piove. 'It is raining.'). Her results suggest that the initial f0 pattern has a very strong effect on the perception of exclamatives (see also Sorianello 2011). One question that has not been addressed in the literature, however, is on what kind of phonological information do listeners rely to identify wh-exclamatives or questions in Cosenza Italian. Consider the example in (1). The prosodic cues that determine the exclamative/interrogative status of sentence (1) could be contained in the Quantifier Phrase ('how many novels'), in the Verb Phrase ('has written') or in the final Noun Phrase ('your friend'). Thus, listeners who rely on the phonological information of one of these constituents, or the combination of more than one, must wait until the relevant material is perceived in order to select the right sentence type between the two available constructions. Another important question that arises from these considerations is whether the critical constituent carrying the phonetic cue or cues that determine sentence type identification is the same for both sentence types and, if it is, whether the the identification of the sentence type is more difficult for exclamatives or interrogatives in real time processing. Our goal is to address these questions with respect to the disambiguation of Cosenza Italian interrogatives and exclamatives such as (1). Our study will shed light on the experimentation in prosody research on the one hand and on processing research on the other.

In order to answer our research questions, we developed an experimental design that involves a speeded identification task with measurement of Reaction Times. This methodology presents some advantages with respect to standard identification tasks in that it allows us to obtain two different behavioral indexes that bring complementary information critical to address the questions discussed above. The off-line judgments may inform us on whether speakers of Italian are able to identify the two sentence types only relying on the phonetic cues. The on-line reaction times will tell us at which point during sentence perception the listeners are able to identify the sentence type and if one sentence type requires more time and cognitive resources for its identification. Finally, the comparison between the off-line and on-line data will allow us to check whether our results converge towards a coherent pattern and whether they are affected by confounds such as speed-accuracy trade off.

The structure of the paper is as follows. Section 2 gives a brief theoretical introduction of the phenomenon including previous investigations. Section 3 presents the methodological part of the paper. In sections 4 and 5, we present the experiment and the results. Section 6 discusses our results and shows where our methodology can be applied in future research.

1. Previous investigations

The prosody of wh-exclamatives has not been studied very extensively so far because wh-exclamatives represent a minor sentence type (see Sadock & Zwicky 1985, König & Siemund 2007). Sorianello (2011) has conducted acoustic experiments for wh-exclamatives and wh-interrogatives in Italian, i.e., the language under investigation here. In her analysis of the Italian variety of Cosenza, different edge tones mark the left edge of the intonation phrases in exclamative and interrogative intonation. In other words, the

initial part of wh-exclamatives and wh-interrogatives would be differentiated by the presence of a high onset in exclamatives (represented by % H, see figure 1 and also Sorianello 2012 for Bari Italian and Avesani 1995, Grice et al. 2005 for Florence Italian), and the missing high onset in interrogatives, see Sorianello 2011):

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

However, it is still a matter of debate if an initial %H boundary tone is appropriate to describe exclamative sentences (see Gili Fivela et al. 2015: 166). Note that the initial boundary tone %H is not a universal feature of exclamatives in general as Gyuris et al. (2013) have shown for Hungarian. According to Sorianello (2011), both sentence types show almost similar endings, i.e. falling intonation at the end of the utterance (see L-L% in figure 1 and 2).[1] The question that arises for the perception of exclamatives and questions is whether the difference in boundary tones at the very beginning of the utterance is what matters for the perception of both sentence types. We will explore this hypothesis in detail below.

Another phonological difference that might play a role for the perception study is the difference in the nuclear accent which is aligned with the stressed syllable of the lexical verb: a low tone L* in exclamatives (figure 1) and low-high tone L+H* in interrogatives (see figure 2):

Figure 2: Information seeking question Che cosa le regalerebbero?' 'What would they give her as a present?' (Gili Fivela 2011 citing Sorianello et al 2011)

The following table shows a summary of the tonal structure in both sentence types according to Sorianello 2011 and Sorianello et al (2011). The most important columns for the perception study are those that represent phonological differences between the two sentence types (highlighted in italics):

	Beginning of the utterance (Initial boundary Tone)	End of the wh-constituent (intermediary tone)	Verbal domain (nuclear accent)	End of the utterance (Final boundary tone)
Exclamatives	*High start (%H)*	Low (L-)	*L**	L-L%

Interrogatives	*No High start (%L)*	Low (L-)	*L+H**	L-L%

Table 1: Description of the tonal structure of exclamatives and interrogatives according to Sorianello (et al) 2011

After having reviewed what we know about intonational structure of both sentence types in (Cosenza) Italian, we will look at perceptional studies on disambiguation between exclamatives and interrogatives.

Gyuris et al. (2013) have run perceptual experiments for Hungarian wh-interrogatives and wh-exclamatives and they found out that Hungarians were better at identifying exclamatives than interrogatives on the basis of intonation. As we will see in our results below, this is different with our Italian participants who did pretty well in judging both sentence types at the end of the utterance in both conditions.

According to our revision of the literature concerning new methodology of prosody (see for example Sudhoff et al 2006, Solé et al 2007), research studies rarely use RT measurements together with judgement tasks for sentence type disambiguation (but see Niebuhr 2007, Chen 2003 for using RTs with respect to gradient and categorical perception of different acoustic signals). By combining the response frequencies with the mean RTs in an identification task, we can investigate the perception of sentence types in a much accurate way than conventional off-line experiments.

Before we present our experiment in section 4, we will discuss different methodologies and argue why we decided to choose the methodology we used for our experiment.

1. Methodology

The goal of our experiment is to see whether Italian listeners from Cosenza can identify exclamatives and questions and if so, if RTs measured in the identification task are the same and whether they differ with respect to different parts of the utterance (e.g. at the beginning or the end of the utterance). In what follows we discuss our research questions in detail and the methodology which will provide answers for these questions.

1. Do Cosenza listeners disambiguate sentence types on the basis of prosody?

2. Do Cosenza listeners identify the sentence type before they hear the end of the utterance and can we localize the potential prosodic cue that triggers their judgement?

3. Can we even identify this cue or cues? For instance, can we test whether an initial or final boundary tone or the nuclear accent is responsible for sentence type disambiguation (see section 2 for Sorianello's acoustic analysis).

4. Can RTs contribute to the study of disambiguation and can we expect different processing of the two sentence types? If listeners identify the sentence type in the same area, we want to know whether the identification task was more difficult in one or the other condition.

In order to test the influence of prosody of exclamatives and questions on sentence type identification of listeners (see research question 1), we employed a judgement task where participants were asked to provide their answer as rapidly as they could. Our main

interest is on whether they could identify the sentence type before they heard the last constituent, and whether they behaved differently in the two conditions (i.e. the question and the exclamative condition). By asking participants to be as fast as possible in their judgements we wanted to trace back their responses to specific parts of intonation (e.g. beginning or the end of the utterance, see section 4 for details) in order to restrict the location of phonetic cues that triggered their response (see research question 2).

In order to localize the cues that triggered the response, it was necessary to make sure that participants respond immediately after they perceived the trigger. If participants respond much later than they have perceived the trigger, we would not be able to correlate their responses with specific locations of potential triggers. We had to find a way to prevent participants from waiting too long to give their response. That is why we asked them to respond as soon as they have identified the sentence type. As will be shown in the results section below, participants did respond as soon as they identified the type, because too early responses that deviated too much from the mean correlated with false judgements. That is to say that those participants that were too fast with their judgements did not provide the right judgements.

Since our goal was to run an experiment for sentence type disambiguation between two types, we gave participants only the choice between two alternatives and did not ask them to formulate their own answer. The latter option is difficult to be combined with RT measurement due to differences in response formulations. We applied an identification/ classification task and not a discrimination task in which participants would have to judge pairs of stimuli as being different. The reason for this decision was that we wanted to focus participants on a particular difference, namely sentence type difference and not on any possible extra linguistic difference (e.g. some emotional or attitudinal difference). The identification task relies on the categorical perception, i.e. the stimuli had to be interpreted categorically as either exclamatives or questions (see also Vanrell 2007, among others for categorical perception of acoustic stimuli). As we did not modify the acoustic signal artificially, we could not test whether the perception of sentence types is a gradient rather than a categorical phenomenon (see Gussenhoven, 2006, Ladd & Morton 1997, Chen 2003 for discussion).

As discussed above, we expected participants to give their responses immediately right after they have identified the sentence type. Our experiment employs thus a similar methodology to the gating paradigm methodology which has been used to study the influence of early phonetic cues in the processing of interrogative meaning (see Petrone 2008 for Neapolitan Italian, among others). In her study, Italian participants had to identify sentence types of spoken stimuli that were gated in specific locations. Her study shows that listeners can recognize sentence types well before the end of the utterance on the basis of prosodic information. This is exactly what we wanted to find out with our participants by presenting them entire natural utterances and not cut or manipulated utterances as usually employed in gating tests (see Sorianello 2012). By presenting entire natural utterances, we tried to test participant's perception in a natural communicational context. We are aware of the fact that by using natural stimuli we cannot control for every variable (e.g. durational difference, see section 6). However, it is the first study of combining responses with RT measurement which we will further develop by modifying

the stimuli in the future (see section 6).

The importance of the reaction time data has been confirmed in a number of studies (Vanrell 2007, Savino & Grice 2008, among others). Therefore, reaction time may provide a more sensitive measure which may reveal differences not shown by an identification task alone. An increase in reaction time would indicate recognition of the crucial acoustic signal. Assuming that phonetic cues have some impact on the perceptual speed for sentence type identification, we should assume shorter RTs in those parts that contain crucial prosodic cues. RTs are longer for ambiguous stimuli and shorter when the stimulus corresponds to an unambiguous category (see Chen 2003). The combination between the processing technique and identification task is not very common methodology in the research on sentence type disambiguation and represents thus a new technique to infer the processing of wh-questions and exclamatives at different locations of the sentence.

In view of the above formulated research questions, we have formulated the following hypotheses:

As our experiment is mainly based on perception of acoustic stimuli, we follow the results of the acoustic analysis of Cosenza Italian wh-exclamatives and wh-questions proposed by Sorianello (2011) (see section 2) and derive our hypotheses about potential acoustic signals that might trigger the disambiguation (see research question 3) on the basis of her acoustic data. However, a noticeable concern of taking her acoustic analysis for granted and apply it to our stimuli has to do with the assumption that some of her descriptions may not apply to our data or may be arguable. We believe however that her analysis is the first step towards building hypotheses and to verify them.

Hypothesis 1: Participants should perceive and identify the sentence type already at the beginning of the sentence, given the difference in the left edge tone (see section 2). As a consequence, differences in the interpretation should be visible when listening to the initial part of the sentence.

Hypothesis 2: Listeners should also identify the sentence type in the verbal region of the sentence because the nuclear accent is different in this region (see section 2). As a consequence, differences in the interpretation should be visible when listening to the middle or verbal part of the sentence.

Hypothesis 3: Listeners will not show any particular difference in responses in the very final end of the utterance because the boundary tone is identical in both sentence types according to Sorianello (2011) (see section 2). As a consequence, differences in the interpretation should not be visible when listening to the final part of the sentence.

As to RT measurement, we would like to formulate the following hypotheses:

We predict equal RTs in regions that have distinct phonetic cues for sentence type disambiguation under both conditions, i.e. the initial boundary tone in the first region and the nuclear tone in the verbal region (Hypothesis 4). This prediction is motivated by the assumption that listeners would have clear cues for sentence type disambiguation under both conditions and should be very secure and thus very fast in their decision (see Pisoni

and Tash 1974 for functions of RTs).

Furthermore, we predict longer RTs in regions that do not show any distinct cues (Hypothesis 5). This prediction is motivated by the absence of any cue which would trigger insecurity of a response under both conditions. We thus predict shorter RTs under both sentence types in the initial part and the nuclear accent part or the verbal part and longer RTs that do not show any phonological distinction, i.e. at the end of the utterance.

We tested these predictions in our experiment (see section 4). The outcome of this experiment is the following. The first hypothesis does not hold for our participants, because there is no significant disambiguation when listening to the initial part. As to the second hypothesis, listeners very rarely disambiguate in the verbal domain. Our results show that the final part does play an important role for sentence type identification (see also Gyuris et al. 2013 on Hungarian). As to RTs, the results show that both sentence types are processed equally long when they gave their responses at the end of the utterance, i.e. no sentence type poses greater difficulties for processing at the end of the utterance. We can thus infer from these measurements that both types represent equally marked sentence types, i.e. one is not more marked or minor (see Sadock & Zwicky 1985, König & Siemund 2007). This is an unexpected result given that exclamatives are used less frequent in corpora of every day speech as our corpus research in C-ORAL-ROM (see Cresti & Moneglia 2005) has shown. It seems that processing does not map necessarily the frequency of use.

1. Experiment

In our experiment, we propose a Reaction Rime (RT) approach combined with identification task to examine the prosodic contrast of two different sentence types. Two variables are measured: (1) response frequencies; (2) mean RTs for identification. Our hypotheses are summarized here again:

1. Participants should perceive and identify the sentence type already at the beginning of the sentence, given the difference in the left edge tone.
2. Listeners should also identify the sentence type in the verbal region of the sentence because the nuclear accent is different in this region.
3. Listeners will not show any particular difference in responses in the very final end of the utterance because the boundary tone is identical in both sentence types according to Sorianello (2011).
4. Both sentence types will trigger shorter RTs for identification in regions that contain phonetic cues for sentence type identification, i.e. the initial domain and the verbal domain.

1. Stimuli

The experimental stimuli were constructed from 20 morpho-syntactically and lexically identical (homophone) wh-exclamatives and wh-interrogatives: 20 sentences with an exclamative and 20 with a question interpretation (see e.g. (1))). For the registration of

the experimental stimuli, we embedded 20 sentences in an exclamative and a question context to elicit the intended sentence type. The stimuli were read aloud by a female Italian from Cosenza. It is important to note that we decided to conduct an experiment recorded by a female speaker involving the most natural prosody possible. We didn't manipulate the length and the F0 of exclamatives artificially because we did not want that an unnatural prosody renders the sentences more difficult to process. Our choice, instead, was to record sentences in a way that they were maximally similar (at the lexical and morpho-syntactic levels), but also maximally natural (at the prosodic level).

All target sentences contained the same syntactic structure: a complex wh-constituent (i.e. quanti,e 'how many' + noun) – a verbal phrase which consists of an auxiliary and past participle – a noun phrase which has the function of a subject which included a definite article (otherwise omitted in the case of immediate family members):

(1) 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!"

As we will show in section 5, the length of the postverbal subject played an important role in RTs of sentence type identification.

We also registered 20 fillers of 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.). The fillers were produced by the same female speaker that produced target sentences. A complete list of experimental target sentences and the fillers is given in the Appendix.

For the perception experiment we used a user-friendly open source script that exploits Pygame (a module for Python). For the evaluation of the data we delimited the target sentences in 4 critical points or "marks" (M1-M4): M1= the beginning of the utterance; M2= end of the wh-constituent (whP) ; M3= end of the verbal constituent (VP) and M4= end of the nominal constituent (NP) after the verbal constituent which was also the end of the utterance (M2-M4 are represented by each closing bracket] and M1 by opening bracket [ ) (see Appendix for information on duration of all marks):

(1) Stimuli with markers

a. [=M 1 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 in three regions. The quantifier region included the first quantifier phrase at the beginning of the untterance (e.g. Quanti romanzi), the verb region (e.g. ha scritto) and the final region included the subject noun phrase at the end of the utterance (e.g. la tua amica). Statistical analysis of the length of the three regions show that the quantifier region was significantly longer in the exclamative condition than in the interrogative condition (excl.= 616 ms vs. inter.= 546 ms, $p<.01$). The verb region was not different in the two conditions (excl.= 592 vs. inter.= 604, $p=.40$). The nominal constituent after the verbal domain was longer in the exclamative condition (inter. = 751 vs. excl.= 933, $p<.001$). As will be shown by our results in section 5, the significant difference in duration of the quantifier region did not have any effect on sentence type identification. The duration seems to be unimportant for sentence type disambiguation according to our experiment.

1. Participants

Participants for the identification task consisted of native speakers of Italian from Cosenza (10 females and 8 males). 15 subjects were university students between 19 and 24 and 3 university workers at age 34. They reported no hearing problems.

1. Tasks and procedure

We ran a sentence type identification experiment in which 18 Italian subjects from Cosenza were asked to press on a button after they have identified the sentence type of non-manipulated utterances. They were asked to press on two buttons on a keyboard as soon as possible with accuracy without guessing. This instruction was made because one of the goals of the experiment was to elicit as many 'early answers' as possible, i.e. whether participants could identify the status of the sentence before the end of the utterance (e.g. just from hearing the quantifier constituent or the verbal constituent and not wait until they hear the last nominal constituent). Before the experiment, specific instructions were given and listeners had a short practice session, with four practice trials for sentence type identification. The experiment lasted about 10 minutes. Decisions made and reaction times (measured from the offset of the stimulus) were stored in the computer. Stimulus presentation and response collection were performed by Pygame (a module for Python).

In the sentence type identification experiment listeners decided about the sentence type of each stimulus they heard by pressing one of the two buttons on the response keyboard. They were not only asked to avoid errors but also to perform the task as quickly as they could manage (time pressure). The two buttons were evenly spaced.

Each trial began with a question to the participant if the participant is ready to start. A beep signaled that an utterance is going to start in order to draw participant’s attention to the stimulus. The session began with the presentation of the auditory stimulus (the target sentence) accompanied by the same sentence visualized in the center of the screen in orthographic transcription and without diacritic marks. The sentence remained on screen until the participant pressed a button. The participants were asked to judge whether the heard sentence was a question or an exclamation. They made their choice by pressing a

button on the keyboard.

For each listener, the experiment included 2 blocks (i.e. the stimuli were divided into two parts). Block 1 contained 10 wh-exclamative sentences and 10 wh-interrogative sentences that were not lexically identical with the exclamative sentences of the same block ("non-minimal pair condition", e.g. 'How many poems your professor has written!' vs. How many cigarettes smokes our father?). 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 contained thus 30 sentences in total. Block 2 contained again 10 wh-exclamatives, 10 wh-interrogatives and 10 fillers. This time, 5 of 10 wh-exclamatives and 5 of the 10 wh-interrogatives were lexically identical with 5 wh-interrogatives and 5 wh-exclamatives already presented in the first block. This is to say that each participant heard the same sentence under exclamative and interrogative condition (i.e. "minimal pair" condition). In total, each participant heard 60 sentences (i.e. 20 exclamatives, 20 interrogatives and 20 fillers). The stimuli were divided into four counterbalanced lists, to which participants were randomly assigned. Listeners were tested one at a time in a quiet room.

1. Statistical analysis and results

The response frequencies for each sentence type and each region of the sentence type are given in Table 2. The response frequency strongly suggests that Italian subjects had no difficulty in interpreting both sentence types. The rejection of the outliers was performed on the trials from the remaining participants (n=17) by excluding every response that occurred during the first region (n=2) or after 6 seconds (about 2.5 standard deviations from the mean of all reaction times) from the end of the utterance (n=5), excluding 1% of the trials (7/680).

1. Off-line results

The overall accuracy was very high in both experimental conditions. Participants correctly identified exclamatives in 93.4% of the trials and interrogatives in 93.7% of the trials. The Generalized Linear Mixed Models conducted on the correctness of the choices did not reveal any significant difference between the exclamative and the interrogative condition (p=.86) as well as between the first and second block (p=.42), i.e. there was no effect on responses with respect to whether the listeners heard lexically identical exclamatives and interrogatives in the same block ("minimal pair condition"):

	Condition exclamatives		Condition interrogatives	
	Correct responses	Incorrect responses	Correct responses	Incorrect responses
Quantifier region M1-M2	0	0	0	0

Verb region M2-M3	1	0	0	0
Final region M3-M4	24	6	26	1
After M4	289	16	290	20

Table 2: responses given in each time region.

Although participants were instructed to make their choice as soon as they identified whether the utterance was exclamative or interrogative, the great majority of responses were provided after the end of the sentence (late responses : 90.7% in the exclamative condition vs. 92.0% in the interrogative condition). In only 31 and 27 trials, in the exclamative vs. interrogative condition respectively, participants provided early responses, which were all made during the listening of the final region except one trial (see Table 2). In the exclamative condition, 25/31 early responses were correct, whereas 6/31 were incorrect. In the interrogative condition 26/27 were correct vs. only 1/27 was incorrect. This could indicate that participants were more prone to error when providing an early answer in judging exlamatives than in judging interrogatives. However, this difference is due to very few trials and in the GLMM including the main factors correctness and prepost (i.e. early vs. late answer) the interaction between these factor is marginally significant (estimate= 2.06, z= 1.77, p= 0.07).

The analysis including only correct answers shows that participants provided a comparable rate of early responses in the two conditions (exclamatives : 9.2% vs. interrogatives : 8.0%). This means that none of the two experimental conditions did show any facilitation effect with respect to the early identification of the status of the utterance as attested by the GLMM conducted on the prepos categorical variable including condition as main factor (p=.99).

1. On-line results on Reaction Times (RTs)

To investigate whether the reaction times obtained from the identification task are affected by the type of the sentence when participants did or did not wait until the entire sentence was heard we conducted a series of analyses that we illustrate in the following paragraphs. Data preparation was obtained by excluding the incorrect answers and by applying a logarithmic transformation to the reaction times. The dependent variable we focus on is response time measured at the end of the sentence, which is positive for the trials in which participants provided a late answer vs. negative for the trials in which they provided an early answer (i.e. before the end of the sentence presentation). After the logarithmic transformation we excluded two outliers presenting a value that was smaller than 3 standard deviations (i.e. more negative than -1). Statistical assessment was achieved by conducting a series of Linear Mixed Models (LMM) on the response times adopting item and participant as random factors with free random slopes and intercepts with respect to the factor of main interest, i.e. condition (cf. Barr et al. 2013).

First, we investigated the effect of the repetition of the test sentences in the second block. From the LMM including the main factors condition and block it emerges that repetition did not have any influence on the response times, as attested by the non-significant main

effect of block (estimate = -0.009, t=-0.228, p=.82) and interaction between condition and block (estimate = 0.057, t=1.084, p=.29). This allows us to include the trials from the first and the second block in the following analyses. Instead, the main effect of condition was significant (estimate = 0.071, t=2.939, p=.009) and it was due to the difference in the response times of the two conditions. In the exclamative condition participants took 525 ms on average to provide a response whereas they took 639 ms in the interrogative condition.

The next analysis focuses on the response times of the two conditions depending on whether participants answered before the offset of the test sentence. When participants provided an early response they pressed the button 259 ms before the end of the sentence in the exclamative condition vs. 239 ms in the interrogative condition. When participants 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 are statistically significant we conducted a LMM adopting condition (exclamative vs. interrogative) and prepost (before vs. after the end of the utterance) as fixed factors and participants and items as random factors, with random slopes and intercepts for condition. The main effect of prepost was significant as expected (p=.000) and the main effect of condition was significant too (estimate = 0.68, t=3.777, p=.001), while the interaction between condition and prepost was not significant (p=.607). Figure 3 presents the response times values for reaction times associated with responses under different conditions, namely when listeners responded before the end and after the end of the utterance:

Figure 3: All log-transformed values of the response times in the four contrasts

Thus, participants were faster at identifying the type of the sentences in the exclamative condition than in the interrogative one regardless of whether they answered before or after hearing the whole sentence. Recall, however, that the last region of the sentence was statistically longer in the interrogative sentences than in the exclamative sentences. As in interrogatives the stress on the final constituent is more prominent, this fact is reflected on our stimuli that were recorded naturally and without any artificial modification of the speech stream. Thus, in order to verify whether the advantage in response time displayed by exlamative sentences was affected by the different length of the Final region in the two conditions we performed a new analysis adding this variable as covariate. The LMM conducted on this analysis was identical in every respect to the previous one, except the inclusion of the length of the Final region as covariate. The results from this model show that the main effects of prepost (p=.000) and length of the Final region (p=.001) are significant, as expected. Critically, however, the main effect of condition was not significant anymore (estimate = 0.013, t=0.611, p=.546) and the interaction between condition and prepost was not significant either (estimate = 0.110, t=1.133, p=.237).

These results suggest that the significant main effects of condition revealed by the previous analyses were mainly caused by the length of the Final region, which is greater in the interrogative sentences than in the exlamative ones. This strongly suggests that in our experiment the type of the sentence had no real effect on the response times regardless of whether the response was provided before or after the sentence presentation.

Although further validation is needed on more subjects, it may be tentatively concluded that both sentence types are equally marked and that this equality is a plausible explanation for the absence of big differences in RTs. Hence, on the whole, the response frequencies and the RTs show no distinction in the identification of interrogatives and exclamatives.

1. Discussion and conclusions

Our experimental design that combined response frequencies with mean RTs in an identification task, has proven a better and more accurate investigation of the perception of sentence types than conventional off-line experiments. RT measurement gives us insights in the speed of processing from which we can infer differences in sentence types on the level of their informational load, processing difficulty and markedness. Our results seem to suggest that both types are equally processed on all of these levels. Neither type is more informationally loaded, more difficult for processing or more marked than the other. This means that both types contain the equal amount of phonological information for sentence type disambiguation, since "reaction time is a monotonically increasing function of the amount of information in the stimulus series" (Hyman, 1953). The equal markedness of both types is surprising given that exclamatives are used less frequent in corpora of every day speech as our corpus research in C-ORAL-ROM has shown. We can thus conclude that the frequency of use of a sentence type does not have any effect on the speed of sentence type recognition.

RT measurement made it possible to observe that the length of the postverbal nominal constituent (see underlined constituent in Quanti romanzi ha scritto <u>la tua amica</u>) affects the rapidity of sentence type identification in a significant way, and it does so under both sentence type conditions. The observation that the last constituent had an effect on RTs suggests that the longer the constituent after the verb is, the more listeners have time to parse the prosodic information crucial for sentence type disambiguation. That is why RTs are shorter with longer constituents in both conditions.

Our study has also shown that the lexical repetition of the stimuli did not have any significant facilitation effect on the response time or the likeliness of anticipating the answer. This result is not very surprising given that lexical information does not contain any information about sentence types. What is even more important for the fact that the lexical repetition did not affect RTs is the observation that participants wait until the end of the utterance in order to disambiguate the sentence type. This suggest that they concentrate much more on the phonological information at the end of the utterance than on the lexical information.

The identification task has shown that listeners do perceive sentence type differences on the level of prosody (see hypothesis 1 in section 3). The fact that listeners gave mostly

correct responses in the great majority of trials and that these responses were not given by chance suggests that there must be a phonological trigger for listener's judgements. In order to find out where this phonological trigger might be, i.e. in which part of the utterance, we subdivided the utterance in different parts and asked the participants to give their responses as to what sentence type they hear as soon as possible. The results show that listeners waited for the end of the last constituent to give their responses (i.e. the nominal constituent at the end of the utterance) in the great majority of trials (more than 90%). This suggests that subjects have to process intonational cues relatively at the end of the sentences for disambiguation contrary to what we would expect under Sorianello's (2011) acoustic analysis (see section 3, hypothesis 2). Recall that under her analysis listeners should perceive sentence type differences much earlier, already at the beginning of the utterance due to the differences in the initial boundary tone (%H in exclamatives and %L in questions) (see table 1, section 2). However, our study shows that Italian listeners do not rely on the prosody of the initial part (i.e. wh-constituent) to judge exclamatives vs. interrogatives. Moreover, under her analysis listeners should perceive sentence type differences also in the verbal domain which is associated to the nuclear accent of the utterance which is different in both sentence types (L* in exclamatives and L+H* in questions, see table 1, section 2). However, as our study has shown almost no one answered before the end of the utterance. This means that participants could not figure out whether the sentence was an exclamative or an interrogative by just listening to the verbal domain of the utterance. It seems that listeners do not perceive nuclear tone differences on the perceptional level (see section 3, hypothesis 3).

From the observation that participants judged at the end of the utterance, we can conclude that neither the initial boundary tone nor the nuclear tone play an important role for the perception of sentence types like exclamatives and questions (see hypothesis 2 in section 3). It is rather the final part that matters for the distinction of both sentence types. Since both sentence types have the same final boundary tone on the level of articulation according to Sorianello (2011), there must be some other phonetic or phonological cue that triggered the responses at the end of the utterance. As the scope of our study was to test the perception of stimuli and not to propose an acoustic analysis, we must find the triggering cue at the end of the utterance in future research.

To sum up, our experimental design allowed us to infer the location of the cue from speaker's responses and to test hypotheses as to which cue could have triggered the responses. These hypotheses were based on Sorianello's description of the intonation. However, we have shown that none of the acoustic differences proposed by her description did trigger listener's responses on the perceptional level. RT measurement was a necessary tool to find out about sentence type processing and the task difficulty.

In our experiment, we worked with natural non-manipulated stimuli that were produced on the basis of specific pragmatic contexts. In future, we will work with controlled (resynthesized) stimuli in order to control for variables like tonal structure and duration and test the prediction about which acoustic item is crucial for sentence type disambiguation and which one triggers relatively fast RTs. We definitely need a more fine grained acoustic analysis in order to evaluate the results. The current experiment does not allow teasing apart which acoustic signal contributed to the relatively well

performed judgements on sentence type disambiguation at the end of the utterance.

The identification task relies on the categorical and not gradient perception. We did not test the perception of any continuum stimuli in order to find out which stimuli are perceived more categorically than the others (see Niebuhr 2007, Chen 2003). We will consider such tests in future research.

One potential problem for measuring RTs in an identification task where participants 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) whereas visual recognition might diminish this delay significantly and might be an alternative task for sentence type recognition (Carreiras, p.c., Carreiras et al. 2014). We will explore this option in future research.

Further tests are needed to control for the reliance that listeners press the button immediately after they perceived the difference and do not wait too long. One possible control test is to force them to respond earlier than the end of the utterance and see at which point forced answers correlate with correct responses. If our experiment is set up in the correct way, we predict that forced answers that occur before the end of the utterance will correlate more often with false answers. We also predict that RTs will be longer with forced too early answers than with forced answers that occur right after the phonetic cue that triggers listener's response.

1. Appendix_Data Set

Item List_Target sentences	M1	M2	M3	M4
Item List_Target sentences of Interrogatives		Placement of markers in interrogatives on the durational level (measured in seconds		
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 seconds)		
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
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?		
Sei spagnolo o portoghese?		
bella?		
Mi daresti il tuo numero?		
Perch_ piangi?		
Hai visto il mio ragazzo?		
Hai 25 anni?		
Se solo sapessi quello che ho fatto per te!		
Sei una stronza!		
Tu sei l'unico uomo per me!		
C'è qualcuno al telefono!		
Non mi dire questo adesso!		
Forse hai ragione!		
Sei una persona speciale!		
Vieni sta sera!		
Guarda 'sto video!		

Sei bellissima!	

1. References

Avesani, Cinzia (1995): ToBIt: un sistema di trascrizione per l'intonazione italiana. Atti delle V Giornate di Studio del Gruppo di Fonetica Sperimentale (A.I.A.), Povo (TN), Italy, 85-98.

BADIP = Banca Dati dell'Italiano Parlato. URL: [http://badip.uni-graz.at/]

Barr Dale et al. (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: Altmann, H. (Hrsg.), Intonationsforschungen. Tübingen. S. 207-221.

Carreiras, Manuel et al. (2014): The What, When, Where, and How of Visual Word Recognition Trends in Cognitive Sciences (TICS), 18(2):90-98.

Chen, Aoju (2003): Reaction time as an indicator to discrete intonational contrasts in English, Proc. EUROSPEECH 2003, Geneva, Switzerland, 97-100.

C-ORAL-ROM = Cresti, E. / Moneglia, M. (eds.) (2005): C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages, Amsterdam/ Philadelphia.

Eady, S. J., & Cooper, W. E. (1986): Speech intonation and focus location in matched statements and questions. Journal of the Acoustical Society of America, 80, 402–415.

Grice, Martine et al. (2005): Strategies for Intonation Labelling across Varieties of Italian. Sun-Ah Jun (ed.), Prosodic Typology. The Phonology of Intonation and Phrasing. Oxford: Oxford University Press. 362-389.

Gili-Fivela, Barbara et al (2015): Intonational phonology of the regional varieties of Italian. S. Frota & P. Prieto Intonation in Romance. Oxford University Press. 140-197.

Gussenhoven, Carlos (2006): Experimental Approaches to Establishing Discreteness of Intonational Contrasts, in Suhoff, S., Lenertovà, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., and Schliesser, J. (eds.), Methods in Empirical Prosody Research. Berlin: Mouton de Gruyter, 321-334.

Gyuris, Beáta et al. (2013): Experimental investigations on the prosody of Hungarian exclamatives. Ms.

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. T. Shopen (ed.), Language Typology and Syntactic Description. Vol. I. (2nd edition) Cambridge: Cambridge University Press. 276 –324.

Ladd, David Robert & Morton, Rachel (1997): The perception of intonational emphasis:

continuous or categorical? Journal of Phonetics 25: 313-342.

Massaro, Dominic W. (1987): Speech perception by ear and eye: a paradigm for psychological inquiry, Hillsdale, NJ: London.

Niebuhr, Oliver (2007): Categorical perception in intonation: a matter of signal dynamics? In Proceedings of the Interspeech 2007, Antwerp, Belgium, 109-112

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

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

Savino, Michelina & Grice, Martine (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. and Zwicky, Arnold M. (1985): Speech Acts Distinctions in Syntax. In Shopen, T. (ed.), Language Typology and Syntactic Description, 155-196. Cambridge: Cambridge University Press.

Solé, Maria-Josep et al. (2007): Experimental approaches to phonology. Oxford University Press.

Sorianello Patricia (2001): Modelli intonativi dell'interrogazione in una varietà di italiano meridionale, Rivista Italiana di Dialettologia vol. 25, 85-108.

Sorianello, Patricia (2011): Aspetti prosodici e pragmatici dell'atto esclamativo. Studi Linguistici e Filologici Online 9, 287-332.

Sorianello, Patricia et al. (2011): L'Italiano parlato a Cosenza. PPt presented at the workshop ToBI Italiano CRIL, Lecce 29/01/2011.

Sorianello, Patricia (2012): A prosodic account of Italian exclamative sentences: a gating test. Ma Q., Ding H. & D. Hirst (eds.), The Proceedings of the 6th International Conference on Speech Prosody, Shanghai (China), May 22-25, vol. I, 298-301, Tongji University Press.

Sudhoff, S. et al (2006) Methods in Empirical Prosody Research. Methods in Empirical Prosody Research. Berlin/New York: de Gruyter (Language, context, and Cognition. 3).

Vanrell, Maria del Mar (2007): A tonal scaling contrast in Majorcan Catalan interrogatives, in G. Elordieta & M. Vigário (eds.), Journal of Portuguese Linguistics (special issue on Prosody of Iberian Languages) 6.1: 147-178.