

KIM A. GROOTHUIS, GIOVANNI LEO, GIUSEPPE MAGISTRO,
CLAUDIA CROCCO

Prosodic features of pragmatic *quello* in Informal Neapolitan Italian: data from real-world speech¹

Due to contact with the local dialects, the informal varieties of Italian spoken in Campania feature a pragmatic use of the demonstrative pronoun *quello*, which can be used to double an argumental DP or as an expletive subject. The present paper investigates the prosodic properties of *quello* in a corpus collected by means of *Pyrato*, a pipeline to extract data from YouTube. 1775 instances of *quello* have been collected, of which 248 were relevant for the analysis. The data show that pragmatic *quello* bears a pitch accent, unlike demonstrative determiner *quello*. Additionally, a prosodic boundary between pragmatic *quello* and the adjacent co-referent DP can be visually and/or auditorily detected, confirming the absence of an intonational contour arching over the two elements. Assuming that the results of our prosodic analysis of *quello* can be extended to its dialect equivalent *chillo*, we formulate a new hypothesis about the origin of pragmatic *chillo/quello*.

Keywords: Neapolitan Italian, Neapolitan, *quello*, pragmatic expletives, real-world speech.

1. Introduction

Campanian dialects, spoken in the southwest of Italy, show a pragmatic use of the third-person pronoun *chillo*, originally a distal demonstrative (*chillo*, *chella*, *chello* ‘that.M/F/N.SG’) (Gaeta, 2014; Ledgeway, 2000: 77–78, 2009: 290–294, 915–922, 2010, 2013: 281–282; Maturi, 2002: 225–228, 2023: 80–81; Radtke, 1997: 83–85; Sornicola, 1996; Vitolo, 2006). Despite the fact that Campanian dialects are null-subject languages, this pronoun seems to function as a pragmatic expletive subject of impersonal predicates, in the sense that it does not seem to have an argumental role within the sentence, while adding an interpretative nuance (1):

- (1) *Chello chiove.*
that.N rain.3SG
‘That’s because/the fact is that it rains.’ (Sornicola, 1996: 336)

Moreover, the same *chillo* is involved in certain topic constructions², where it can double an expressed subject (2):

¹ This paper is the result of a close collaboration of all four authors; for official academic purposes only, Kim Groothuis takes responsibility for § 2-3, Giuseppe Magistro for § 4, Giovanni Leo for § 5, and Claudia Crocco for § 6. We assume shared responsibility for § 1 and § 7.

² *Chillo* can also double focused subjects, cf. Groothuis (2022) for discussion.

- (2) *Chillo pateme nun vò capi.*
 that.M.SG father=mine NEG want.3SG understand.INF
 ‘My father does not want to understand.’ (Sornicola, 1996: 326)

In both (1) and (2), *chillo/chella/chello* is an optional element in the clause in the sense that its absence does not render the clause ungrammatical, unlike canonical expletives such as ‘there’ or ‘it’ in non-null-subject languages like English. It does not affect the truth value of a clause either, which means that it serves a discourse function, although it is not easy to pin down which one exactly.

Due to contact, a similar pragmatic use of the demonstrative *quello* (*quello/a*, ‘that.M/F.SG’, *quelli/e* ‘those.M/F.PL’; expletive: (3); topic: (4)) percolates into markedly informal varieties of Campanian Italian (De Blasi and Fanciullo, 2002: 636; Maturi, 2002: 225–228, 2023: 139):

- (3) *Quello piove.*
 that.M rain.3SG
 ‘That’s because/the fact is that it rains.’ (De Blasi and Fanciullo, 2002: 636)
- (4) *E certo, quello il brodo di gallina è sostanzioso.*
 and sure that the broth of chicken be.3SG substantial
 ‘And certainly, chicken broth is substantial.’ (De Filippo, Natale in Casa Cupiello)

While pragmatic *chillo/quello* has received some attention within the syntactic literature, its prosodic properties are practically uncharted, in spite of their most likely relevance for syntactic and pragmatic analysis. The present study sets out to contribute to filling this gap by exploring the prosodic properties of pragmatic *quello* in one of these varieties, i.e. Informal Neapolitan Italian (henceforth INI). The paper is structured as follows: § 2 reviews past literature about the intonation of *quello*; § 3 discusses the main aims and research questions; in § 4 we explain the methodology used for collecting data from real-world speech as found on YouTube; § 5 presents the results; § 6 provides a preliminary analysis of the data, and § 7 concludes the paper.

2. The intonation of *quello*: state of the art

The syntactic literature on pragmatic *quello* makes reference to the intonational pattern of the sentences in which it occurs, but no detailed studies on the matter are available. For instance, it has been claimed that *chillo* in utterances such as (2), when it introduces a new or shifted topic, has “the intonational contour [...] of canonical topic-comment structure and not, for example, that of a rhematic sentence” (Ledgeway, 2010: 262). However, the prosodic properties of such a canonical topic-comment form are not explored in the study. Other contexts with *chillo*, specifically copular clauses as (5), are claimed to be often marked by an exclamative contour (Ledgeway, 2009: 291, 294; Sornicola, 1996: 337; Vitolo, 2006: 65). Also in this

case, however, the prosodic properties of *chillo* and the contour in which it appears are not further explored.

- (5) *Chillo* *è* *scemo!*
 that.M.SG be.3SG stupid
 ‘He’s stupid!’

Moreover, *chillo* and the following determiner phrase (henceforth: DP)³ would not belong to the same prosodic phrase, which is taken as evidence that the two do not form one syntactic constituent (Ledgeway, 2009: 815, 2010: 260; Sornicola, 1996: 326–327). Similarly, Maturi (2002: 226) states that “*anche quando quello sia immediatamente adiacente al SN cui è accordato, [...], una breve pausa o una caduta del contorno intonazionale lo separano dal SN, circostanza che porta ad escludere che in questi casi possa essere analizzato come aggettivo*” (also when *quello* is immediately adjacent to the NP with which it agrees, [...], a short pause or a fall of the intonational contour separates it from the NP, something that leads us to exclude that in these cases it can be analysed as an adjective). Such intriguing observations on a prosodic break (*pausa* or *caduta*) between *chillo/quello* and the DP, or the intonational contour of the sentences in which it occurs, however, are regrettably only based on impressionistic evaluations.

Finally, in the literature the idea has been put forward that prosody plays an important role in the distinction between the different pragmatic uses of *chillo/quello*, cf. e.g., Sornicola (1996: 338) who suggests that “*i fattori prosodici abbiano un qualche influsso sull’intera casistica descritta*” (“prosodic factors have some type of influence on the typology as described”). This same idea is repeated by Vitolo (2006: 69). A systematic study of the interaction between prosody and the syntactic-pragmatic uses of *quello* is, however, still lacking.

In this paper, we aim to take a first step at investigating the prosody of pragmatic pronouns in Southern Italy. The data of choice for a first exploration of a phenomenon typically occurring in dialects or colloquial regional Italian are represented by informal, real-world dialogical productions. At present, we do not dispose of corpora of Neapolitan dialect and Neapolitan Italian featuring informal speech with these characteristics. In fact, speech elicited through task-oriented techniques – such as the CLIPS corpus (www.clips.unina.it), which also contains dialogues from the area of Naples – is not suitable for the exploration we wish to conduct, because pragmatic *quello* occurs only seldom in such a speech style⁴. This suggests that the speaking style of such dialogues is not sufficiently informal to encourage the use of a trait which seems, on the other hand, widespread (Ledgeway,

³ The determiner phrase is the constituent formed by a D(eterminer) head (e.g. an article or a demonstrative) and an N(oun)Phrase.

⁴ Pragmatic *quello* occurs only sporadically in the corpus CLIPS/Naples, despite the strong interactivity of the dialogues included in this corpus. Similarly, a preliminary search in a corpus of interviews with university students from Naples (courtesy of Margherita di Salvo) yielded a total of 23 occurrences of pragmatic *quello* in 12 hours of speech.

2010: 259), and presumably not socially marked in Neapolitan Italian (Groothuis *et al.* in prep.). For this reason, we opted for a different approach to the data collection.

The data examined in this work are extracted from YouTube by means of a pipeline relying on the automatic-generated subtitles (see § 4 for a detailed description of the pipeline). Although social media speech is not necessarily spontaneous in the sense of non-read or non-scripted (Beckman, 1997; Tucker and Mukai, 2023; Warner, 2012), it is normally meant to be addressed to a real audience. Such real-world speech, therefore, can be part of the real-life linguistic experience of the speakers from the area of Naples.

At present, data scraping is only possible for Italian since automatic subtitles in non-standard languages such as Neapolitan are currently not supported by YouTube. Accordingly, the present study focuses on *quello* in the INI as spoken in Naples and the surrounding area. While this choice clearly represents a limitation of the study, it is worth stressing that the analysis of *quello* in INI will also be relevant for future studies on the prosody of *chillo* in Neapolitan. Although a dialect and the regional variety of Italian of the same area are two different languages, it is known from past studies that dialects highly influence the Standard variety on all linguistic levels, also as far as prosody is concerned (Barone, in prep; Cataldo, 2023; Crocco, Gili Fivela, and D'Imperio, 2022; O'Rourke, 2005; Telmon, 1994). So, we assume as a working hypothesis that *quello* and *chillo* also resemble each other intonationally. We leave the study of the intonational of *chillo* in Neapolitan for further research (cf. Groothuis & Crocco 2023, in prep.).

3. Aims

The main goal of the study is, thus, to investigate the interaction between the prosody, the syntax, and the pragmatic roles of pronouns such as *quello* in Campania by exploring the prosodic properties of *quello* in a corpus of INI scraped from the Internet. In doing so, we also aim at elaborating the above-mentioned observations originally made on Neapolitan (cf. § 2) concerning:

- a. the presence/absence of an overarching contour including pragmatic *quello* and the adjacent co-referent DP (cf. ex. (4)) (RQ1);
- b. the presence/absence of detectable prosodic differences when pragmatic *quello* is co-referent with a focalized or topicalized constituent (RQ2);
- c. moreover, we intend to compare the intonational contour of the pragmatic uses of *quello* and its use as a non-deictic subject pronoun (RQ3).

The study thereby contributes to the description of the prosody of the Italian linguistic space and especially to the study of informal regional varieties. Since pragmatic *quello* is involved in the expression of topic and focus, the study of its prosodic properties will also contribute to our understanding of the syntax-prosody interface in Italo-Romance.

4. Methodology

To collect a corpus of INI, we relied on the *Pyrlato* pipeline (developed for parallel work, cf. Magistro, Crocco, 2023), written in Python, to extract natural data from YouTube videos, mainly movies, TV shows, and comedy sketches. The corpus therefore includes scripted and spontaneous productions in INI.

First, *Pyrlato* uses the module *PyTube* to search for specific keywords (in our case, names of local shows and public figures)⁵, resulting in a list of relevant videos. *Pyrlato* then checks that the videos are in the desired language, so as to look for specific strings in the language under scrutiny. The pipeline works for Arabic, Dutch, English, French, German, Hindi, Indonesian, Italian, Japanese, Korean, Polish, Portuguese, Romanian, Russian, Spanish, Thai, Turkish, Ukrainian, and Vietnamese. Naturally, we selected Italian for our case study. Next, *Pyrlato* generates an object scraping YouTube's automatic subtitles. These objects contain a Python dictionary with timestamps and the captions of the videos. The captions can be searched to match an exact string or a regular expression. This allows the software to look for more complex patterns such as specific word structures. In our case, a list of the morphological realizations of *quello* (*quello/a/i/e*) was iteratively searched in the captions. If the desired expression is matched, then the respective timestamp is extracted to cut the occurrence from the audio of the video. In our application, the sentence containing the word is extracted in .mp4 and trimmed observing a time-window of 7 seconds. Unfortunately, the timestamps provided by YouTube are not aligned word-by-word but are generally placed between sentential pauses: a cutting window of 7 seconds is then big enough to ensure that *quell** is present in the cut fragment. Finally, .mp4 videos are converted to .mp3 files usable in *Praat* (Boersma, 2001; Boersma and Weenink, 2021).

Data such as those extracted from YouTube are in lossy formats and/or recorded in noisy conditions. For this reason, they are usually discarded in prosodic studies as they do not meet the sound quality standards expected for most acoustic analyses. Such real-world data, however, represent an invaluable source of information concerning the speakers' actual linguistic uses, exactly for their natural character and in spite of their lesser acoustic quality. Of course, this choice has consequences for the type of analyses that can be carried out on the data. In this paper, we focus on the analysis of pitch, which is a relatively robust prosodic feature (Fuchs and Maxwell, 2016). Through auditory and visual inspection in *Praat* (Boersma, 2001; Boersma and Weenink, 2021), we aim at detecting the presence and (if possible) the type of pitch accent occurring on *quello* and on the co-referent DP.

⁵ The following keywords have been used to search for videos in the Italian of Naples and the surrounding area: 'Rita De Crescenzo', 'De Filippo', 'Nino D'Angelo', 'Vincenzo Salemme', 'Made In Sud', 'The Jackal', 'Totò film', 'Canale 21', 'Napoli calcio', 'Cucina napoletana', 'Peppe Iodice', 'Cannavacciuolo', 'Massimo Troisi', 'Simone Schettino', 'Teleluna', 'Telegaribaldi', 'Oscar Di Maio', 'Alessandro Siani', 'Gigi D'Alessio', 'Telecapri', 'Arteteca', 'Biagio Izzo', 'Io speriamo me la cavo', and 'L'amica geniale'.

5. Results

5.1 Observed types of *quello* in the corpus

In total, 1397 videos (ca. 550h) were scraped, in which 1775 instances of *quello* were found. First, all instances of *quello* have been checked for relevance: utterances produced by speakers from regions other than Campania, and purely deictic uses as a demonstrative adjective or pronoun have not been included in the analysis. A sample of 20 randomly selected instances of *quello* as a demonstrative adjective has been considered in a second phase of the research to compare their characteristics with those of pragmatic *quello*. Finally, sentences in dialect (N=14) have been discarded.

All relevant instances of pragmatic *quello* (N=248) have subsequently been categorized into seven types based on the syntactic context in which they occur. Type I (6) consists of a *quello* reduplicating a preverbal subject (6), whereas type II reduplicates a postverbal subject (7). Types III and IV are pronominal uses of *quello* as the subject of a non-copular (8) and copular verb (9), respectively. Type V refers to the use of *quello* in a cleft sentence (10), where it occurs in initial position, preceding the copular and the clefted constituent, with which it is co-referent (10). In type VI, *quello* is co-referent with a focused DP preceding a copular verb (11) (11). Finally, type VII is the use of *quello* as an expletive without any co-referent DP (12) (12).

- | | | | | |
|------|---|----------------------|---|--|
| (6) | <i>Quella</i> | <i>mammà</i> | <i>sta</i> | <i>in paradiso.</i> (<i>La lettera di mammà</i>) |
| | that.F.SG | mom | be.3SG | in paradise |
| | 'Mom is in heaven.' | | | |
| (7) | <i>Quelli</i> | <i>rimangono</i> | <i>impressionati</i> | <i>i bambini.</i> (<i>Bello di papà</i>) |
| | that.M.PL | remain.3PL | impressed | the children |
| | 'The children get creeped out.' | | | |
| (8) | <i>Quella</i> | <i>a te</i> | <i>ti ascolta.</i> (<i>Giuro che ti amo</i>) | |
| | that.F.SG | to you | you=listen.3SG | |
| | 'She listens to YOU.' | | | |
| (9) | <i>E quell'</i> | <i>è</i> | <i>l'umidità.</i> (<i>Totò, Peppino e le fanatiche</i>) | |
| | and that.F.SG | be.3SG | the=humidity | |
| | 'And that's (because of) the humidity.' | | | |
| (10) | <i>Quello</i> | <i>è lui ... fra</i> | <i>Martino... che</i> (<i>I 4 monaci</i>) | |
| | that.M.SG | be.3SG him friar | Martino that | |
| | 'That's him... friar Martino... who...' | | | |
| (11) | <i>Quelle</i> | <i>due</i> | <i>so'.</i> (<i>Arteteca</i>) | |
| | that.F.PL | two | be.3PL | |
| | 'They're TWO.' | | | |
| (12) | <i>Quello</i> | <i>eravamo</i> | <i>prima un gruppo più ampio.</i> (TGR) | |
| | that.M.SG | be.IPFV.1PL | first a group more large | |
| | 'At first we were a large group.' | | | |

The number of found instances per type is shown in Table 1, column "N".

Next, the instances of *quello* have been inspected visually and auditorily using *Praat* (Boersma, 2001; Boersma and Weenink, 2021). During this stage, we examined the pitch contour on the entire utterances, i.e., by checking both *quello*

and – if present – the co-referent DP. Some instances could not be processed because the audio quality was too low to appreciate reliable pitch contours ($N = 53$). This was anticipated given the nature of the data (cf. discussion in § 4). To assess the presence of a pitch accent, we relied both on the pitch track observation and on the auditory appreciation of a prominence on *quello*. Note that the unvoiced onset in the stressed syllable of *quello* [ˈkwel] can make the visual appreciation of the pitch accent more difficult. Beside creating a gap in the pitch track, voiceless obstruents also raise f_0 in the first part of a following vowel in several languages including Italian (Hanson, 2009; Kirby and Ladd, 2016). However, coherently with the regional features characterizing Campania varieties, in our data voiceless stops are often realized as voiced.

Table 1 - *Pitch-accented instances of quello per syntactic type*

SYNTACTIC TYPE	N	PA	NON-PA	?	LOW-Q
I: <i>Quello</i> _i DP _i Pred	30	16	1	5	8
II: <i>Quello</i> _i Pred DP _i	6	5	0	1	0
III: <i>Quello</i> _{subj}	99	62	2	15	20
IV: <i>Quello</i> _i Cop DP _i	102	61	5	11	25
V: <i>Quello</i> _i cleft DP _i	6	4	0	2	0
VI: <i>Quello</i> _i DP _i Cop	1	1	0	0	0
VII: <i>Quello</i> _{Expl}	4	3	0	1	0
Total	248	151	8	35	53

5.1.1 *Quello* as weak deictic subject pronoun

Before we proceed to the presentation of the results, a clarification is needed concerning the inclusion of *quello* as a weak deictic subject pronoun (Type III and IV, examples 8-9) in the corpus⁶.

The use of *quello* as a weakly-deictic subject pronoun may seem *prima facie* unrelated to the other cases examined in this study for two reasons: first, the cases included in the analysis seem identical to other possible instances of *quello* being used as a deictic pronominal subject in standard Italian (e.g. “*trovi brutto anche il cappello?*” “*no, quello mi piace*”; “*quella è la casa dei miei sogni*”); second, they look different from the other pragmatic cases where *quello* is co-referent with a DP belonging to the same sentence (e.g. “*quello tuo figlio sta sempre da solo*”). However, it can be argued that the cases of *quello* as a weakly-deictic subject pronoun included in the analysis exhibit a specific pragmatic nuance that justifies their inclusion in the corpus. In this subsection, we will briefly explain why we think this use of *quello* is pragmatically marked.

⁶ By defining such cases as weakly-deictic, we wish to stress that, in many cases, although the pronoun does not have a clear deictic function, it can be argued that pronoun has some kind of discourse reference. In fact, the distinction between deictic and non-deictic *quello* is often hard to make.

In cases such as “*quella a te ti ascolta*”, *quella* has a weak deictic function, as it simply refers to an identifiable subject. In a less regionally flavoured version of the same sentence, *quella* could be replaced by *lei* (“she”), without changing the meaning of the sentence: “*lei a te ti ascolta*”. The use of *quello/quella* as a subject pronoun is likely to be a regional use rooted in the dialect: the expansion in 18th century of *chillo/chella* in Neapolitan as a subject pronoun at the expense of the subject pronoun *isso/essa* (i.e. without a clear deictic function) has been pointed out by Ledgeway (2009: 276-280, 286-290). Similarly, the Neapolitan speakers of Italian can use both *quello/quella* and *lui/lei* as subject pronouns (both sentences “*quella a te ti ascolta*” and “*lei a te ti ascolta*” are acceptable in the variety). While the alternation between *quello/quella* and *lui/lei* can possibly be affected by the speech style – regional features tend indeed to emerge in informal registers (Berruto, 1987 [2012]) – it can be argued that the expansion of *quello* as a subject pronoun is pragmatically motivated.

Moreover, the two variants of the subject pronoun (*quello/quella* vs *lui/lei*) are not freely interchangeable but have different syntactic properties, as in Neapolitan (Ledgeway, 2009: 290). In particular, two syntactic differences can be noted. First, only *isso/essa* can have a local antecedent (i.e., be co-referent with a DP in the same clause), cf. (13) where only *essa* is grammatical. Second, in an embedded clause, *isso* can refer anaphorically to the main clause subject or have a disjunctive interpretation; *chillo*, on the other hand, can only be interpreted disjunctively (see example (14)):

- (13) *Vicenza_i s'è messa a vnnere 'o ccaffè pur'*
 Vicenza REFL=be.3SG put to sell.INF the coffee also
*essa/*chella_i*
 she/that.F.SG
 ‘Vicenza started to sell coffee also herself.’ (Ledgeway, 2009: 920)
- (14) *Ciro_i ha itto ca isso_{i+j}/ chillo_{i/j} non capesce*
 Ciro have.3SG said that he that.M.SG NEG understand.3SG
niente.
 nothing
 ‘Ciro has said that he does not understand anything.’ (Ledgeway, 2009: 920)

The same can be noticed for regional Italian *quello*, which differs from *lui/lei* along the same lines:

- (15) *Vincenza s'è messa a vendere il caffè pure*
 Vicenza REFL=be.3SG put to sell.INF the coffee also
*lei /*quella.*
 she /that.F.SG
 ‘Vicenza started to sell coffee also herself.’
- (16) *Ciro ha detto che lui_{i+j}/ quello_{i/j} non capisce*
 Ciro have.3SG said that he that.M.SG NEG understand.3SG
niente.
 nothing
 ‘Ciro has said that he does not understand anything.’

Also more generally, the use of a subject pronoun in a null subject language is automatically marked. In unmarked situations, the more natural choice would be to use no pronoun (cf. the Avoid pronoun principle, Chomsky, 1981). A subject pronoun is used when it has a marked pragmatic interpretation (be it topic or focus). This is in line with the Division of labour principle as proposed by Horn (1984), according to which the use of a marked expression when a corresponding unmarked form is available tends to be interpreted as conveying a marked message.

In other words, speakers can use *quello* as a subject pronoun to indicate that the subject has a degree of topicality, as it is present in the discourse, though not necessarily in the immediate context in which the interaction takes place. This is unsurprising given the high degree of topicality of pronouns and the frequent overlap of the categories topic and subject (Chafe, 1976). Therefore, the case of *quello* as a subject pronoun shows that, in certain contexts, the boundaries between an unmarked pronoun and a topicality marker are blurred.

5.2 Prosodic inspection of *quello* per type

From the observation of our corpus, *quello* appears to be marked by a high or rising f_0 movement that can be interpreted as a pitch accent in 152 out of 248 occurrences across the different syntactic-pragmatic contexts. In other cases, the presence of a pitch accent could not be determined for sure (N=35). The results of this inspection can be seen in Table 1, where PA = presence of a pitch accent on *quello*, NON-PA = absence of pitch accent, ? = uncertainty about the presence of the pitch accent and LOW-Q indicates those data excluded in the analysis because of the low quality of the audio. In the remainder of this section, we will discuss the results more in detail per type of *quello*.

When occurring with a co-referent DP (Types I and II), both *quello* and the co-referent DP seem to carry a pitch accent (cf. Fig. 1, 2, 3 and 4). Out of 36 instances, a pitch accent on *quello* could be individuated in 21 cases. There were 14 cases where the presence of a pitch accent was uncertain or could not be determined due to low signal quality. Only in one example, there was no pitch accent on *quello*.

In the cases where *quello* and the co-referent DP are both preverbal (Type I), they mostly belong to the same major prosodic phrase, although a degree of disjuncture between *quello* and the DP can be detected, hinting to a minor prosodic break separating the two elements. Additionally, *quello* and the co-referent DP can be separated by a pause (Fig. 4), a pitch drop (Fig. 3) or by intervening material (Fig. 2). Such cases strengthen the observation that *quello* and the co-referent DP are not necessarily included in the same phrase.

Figure 1 - *Example of quello + preverbal DP (Type I, “That Mom is in paradise”)*

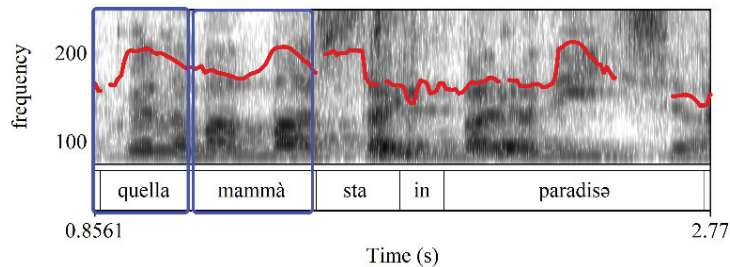


Figure 2 - *Example of quello + preverbal DP with intervening material (Type I, “That sometimes loneliness plays nasty tricks on you”)*

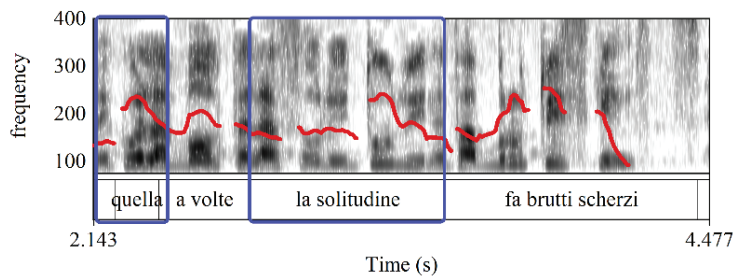


Figure 3 - *Example of quello + preverbal DP with pitch drop (Type I, “That uncle Pasquale has just arrived from Switzerland”)*

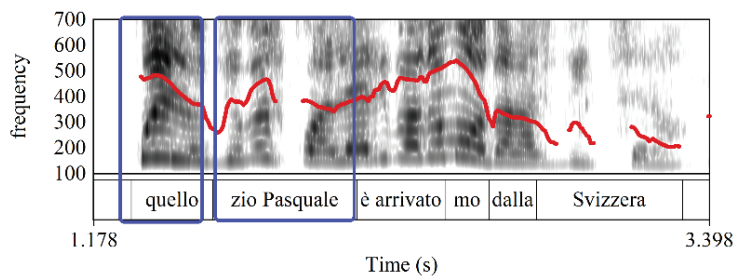
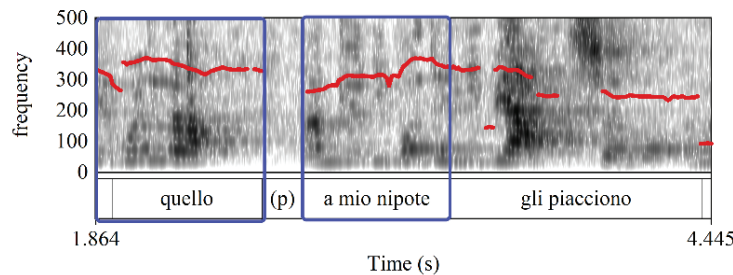


Figure 4 - *Example of quello + preverbal DP with pause, indicated in transcription as “p” (Type I, “That my nephew likes them”)*



For type II, where the co-referent DP is postverbal, a pitch accent can mostly be detected on *quello* (see Fig. 5 and 6). The co-referent DP can be in post-focal position (Fig. 5).

Figure 5 - Example of *quello* + postverbal DP
(Type II, “Those they get creeped out, the children”)

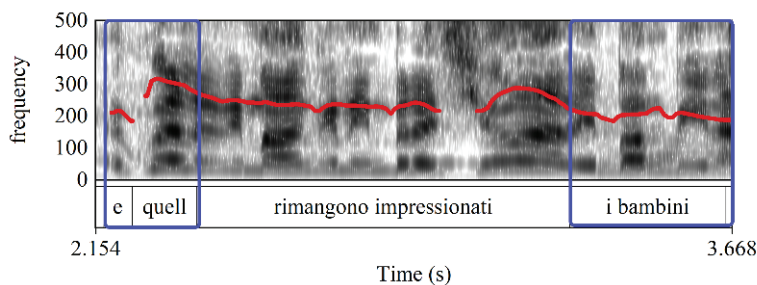
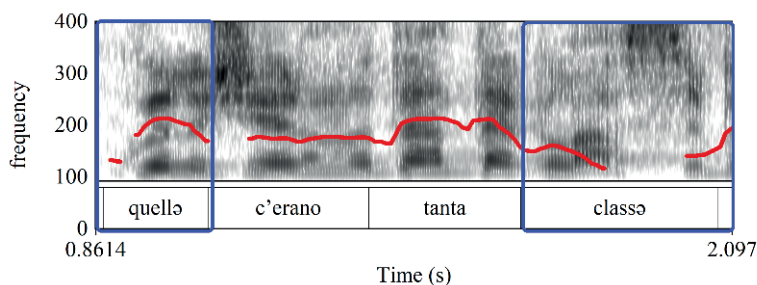


Figure 6 - Example of *quello* + postverbal DP (Type II, “Those there were many classes.”)



Overall, in those cases where *quello* is associated with a co-referential DP, both items seem to be pitch accented, although the properties of the accented DP depend on its focal or post-focal position, as expected in this variety (D'Imperio 2001). The case presented in Fig. 5 illustrates a post-focal DP *i bambini* (“the children”), functioning as sentence topic, which is realized with a reduced or downstepped pitch movement (D'Imperio, 2001). In Fig. 6 the DP *tanta classe* (“so many classes/classrooms”) is focused and carries the nuclear accent.

When used as a subject pronoun, both of copular and non-copular verbs (types III and IV), *quello* carries a pitch accent of its own, as shown in Fig. 7 and 8. The shape of f_0 movement is similar to the one we observed in the previous types, i.e., a rise / high.

Figure 7 - Example of *quello* as a subject pronoun DP, non-copular verb
(Type III, “That he will arrest us”)

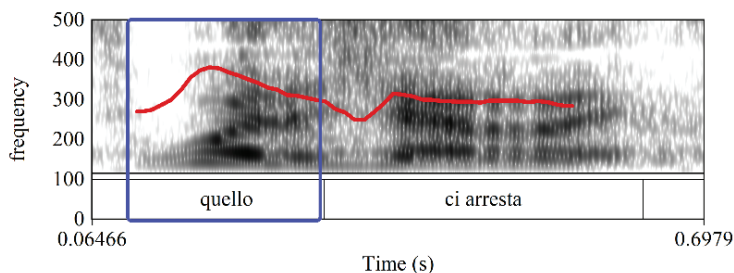
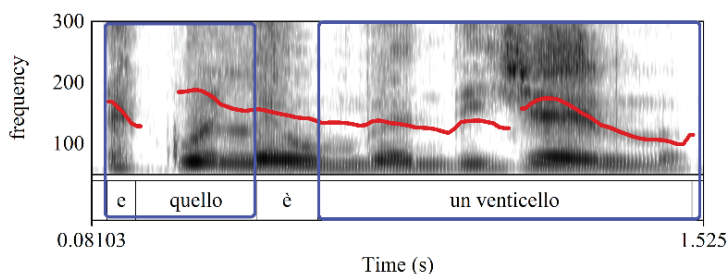


Figure 8 - Example of *quello* as a subject pronoun DP, copular verb
(Type IV, “That it is a little breeze”)



Also in types V and VI, *quello* exhibits a pitch accent. We can see this in Fig. 9 for type V, where a focused DP occurs before the copula, and in Fig. 10 for type VI, where the focused DP is a clefted constituent.

Like the other types, expletive *quello* is also marked by a high/rising pitch accent. An example can be seen in Fig. 11. In this case, *quello* is not co-referent with an argument within the sentences but refers to a proposition present in the discourse (cf. Ledgeway, 2010: 286-288; Groothuis, Crocco, in prep.).

Figure 9 - Example of *quello* with a focused DP (Type V, “Those they are TWO”)

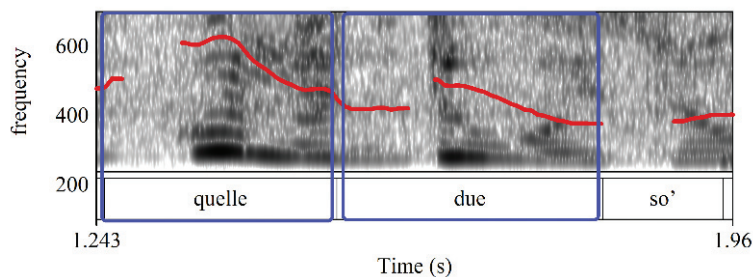


Figure 10 - Example of *quello* with a clefted DP
(Type VI, “That it is the director let’s say that tends to...”)

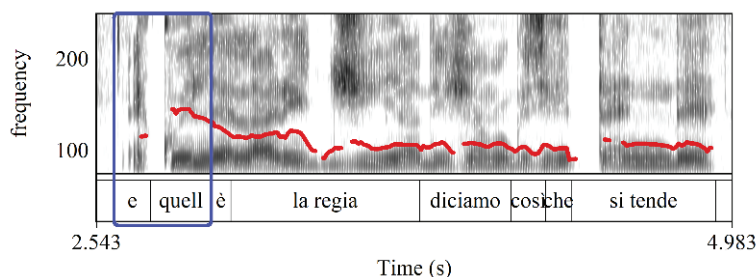
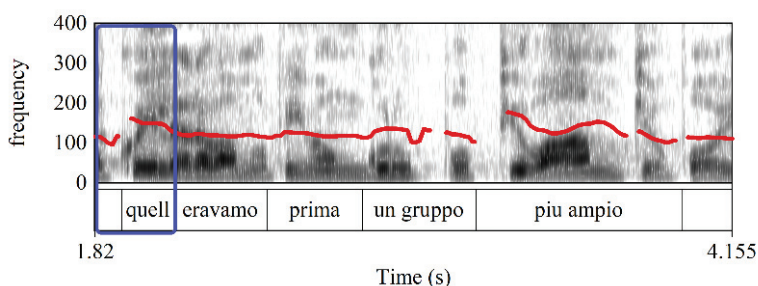


Figure 11 - Example of expletive *quello* (Type VII, “That before we were a bigger group”)



As far as can be observed from the corpus, no detectable prosodic differences emerge in the realization of *quello* according to the pragmatic context in which the pronoun appears. *Quello* can indeed refer to a topical or focal DP, but the discourse category of the co-referent DP does not seem to affect its realization. However, given the low acoustic quality of the data at our disposal, more targeted investigations are necessary to confirm this observation. At the same time, the accent marking the DP does vary according to its (pre-, post-, or focal) position in the utterance, as expected based on previous evidence on Neapolitan Italian (D’Imperio, 2002).

6. Discussion: demonstrative *quello* vs pragmatic *quello*

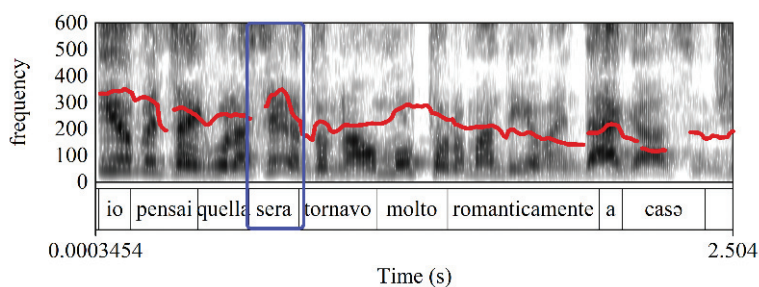
Through *Pyrlato*, we were able to extract ca. 200 occurrences of pragmatic *quello* that could undergo a prosodic analysis out of a total amount of 1729 items. While this may seem a very small amount of usable data, the picture changes if we consider that spontaneous utterances containing a pragmatic *quello* are difficult to obtain⁷. With this in mind, automatic scraping of internet data represents a valuable technique to gather informal data suitable for the investigation of phenomena that are otherwise difficult to observe.

⁷ Cf. fn. 4.

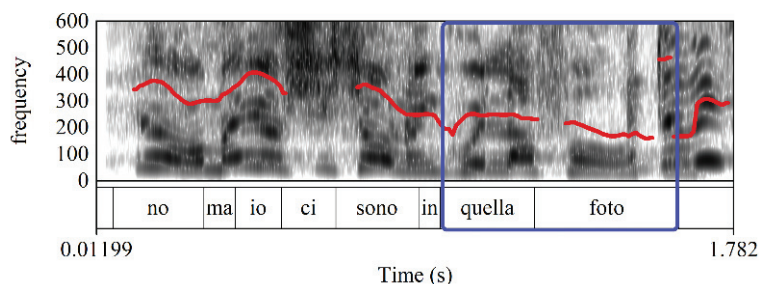
With regard to the first research question (RQ1) about the presence/absence of an overarching contour including pragmatic *quello* and the adjacent co-referent DP (Type I), the data indicate that both elements are pitch-accented and can be separated by an audible boundary. However, the strength and the actual realization of such a boundary, e.g., as a tonal event, can vary in our data. Occasionally, *quello* and the following DP can be separated by intervening material (Fig. 2) as well as by a pause (Fig. 4).

Such observations concerning Type I *quello* can be integrated with the results of a comparison with the context in which *quello* appears as a demonstrative determiner. In many of the occurrences, collected by means of *Pyrlato*, *quello* is a demonstrative determiner (i.e., *quello* with an adjectival value, as in *quella casa* “that house” or *quel ragazzo* “that boy”), as expected in Standard Italian. We inspected a sample of 20 cases of *quello* used as a demonstrative determiner. In these instances, the modified noun rather than the pronoun *quello* seems to be the bearer of the pitch accent (Fig. 12 and 13)⁸. Furthermore, no audible boundary can be detected between *quello* and the modified noun, that seem therefore included in the same prosodic phrase. With respect to Type I *quello*, demonstrative determiners seem to feature a different prosodic organization, characterized by the “overarching contour” hypothesized a.o. by Ledgeway (2010). By and large, our results for Type I *quello* converge with the observations put forward by Sornicola (1996), Maturi (2002), and Ledgeway (2010) on Neapolitan *chillo*. This result also provides support to the related hypothesis that the two items do not form one syntactic constituent, unlike determiner *quello* and the modified noun.

Figure 12 - Example of demonstrative determiner *quello*
 (“I thought that that night I would go back home very romantically”)



⁸ Only in one case (*quelle caratteristiche*, ‘those features’) both *quello* and the following noun are clearly accented.

Figure 13 - Example of demonstrative determiner *quello* (“No, but I am there in that picture”)

The data gathered through *Pyrlato* indicate that the prosodic properties of the co-referential DP vary according to its position with respect to the focus, thereby confirming the findings of previous research on the Neapolitan variety of Italian (D’Imperio, 2001, 2002). As far as it is possible to observe in the data examined in this paper, *quello* does not present detectable prosodic differences when it is co-referent with a topicalized or focalized constituent (RQ2). However, we are not able to further elaborate on this point due to the signal quality of the collected material. Therefore, as the prosodic realization of *quello* cannot be adequately explored in the available dataset, we leave this issue open for a different work based on more controlled data (Author, Author in prep.).

We also intended to compare the intonational contour of the pragmatic uses of *quello* with its use as a weakly-deictic subject pronoun (RQ3). As discussed in § 5.1.1, the boundaries between an unmarked pronoun and a topicality marker are blurred in the case of non-deictic pronominal *quello* (Type III-IV)⁹. Also at the prosodic level, we do not see a clear difference between *quello* as a weak deictic subject pronoun and the pragmatic uses of *quello*. In fact, we notice that *quello* in type I-II, which has a clear pragmatic function, and *quello* in types III-IV, in which it could be ambiguously interpreted as a topic and a subject, feature a high/rising pitch accent. No further prosodic property is found to have a distinctive role across these categories. Instead, the relevant prosodic distinction seems to be between the demonstrative determiner *quello* (as in *quella sera*, “that night”) and the other uses examined in this study: while the demonstrative determiner does not normally bear a pitch accent, it tends to be pitch accented when it occurs with a pragmatic or pronominal value.

On the assumption that the results of our prosodic analysis of *quello* can be extended to its dialectal equivalent *chillo*, we formulate a new hypothesis about the origin of the pragmatic use of *chillo/quello*, namely that the pitch-accented demonstrative pronoun *chillo/quello* may have provided prosodic cues for the functional reanalysis of pronominal *quello* as a pragmatic marker. When used as a subject pronoun, *quello* must bear a certain degree of discourse salience, because

⁹ Note, however, that *quello* can be co-referent with a focused element (type V and VI), suggesting that this pronoun does not mark topicality per se.

otherwise a null pronoun/pro-drop would be preferred (see § 5.1). Because of its salience at the level of the discourse, the pronoun interpretation is ambiguous between a regular preverbal subject pronoun in the IP-domain and a pragmatic category (i.e., a topic or a focus element) encoded in the left periphery of the clause (Ledgeway, 2010; Rizzi, 1997; Rizzi and Bocci, 2017). We propose that it is exactly this ambiguous context which has permitted the reanalysis of *quello/chillo* as a left-peripheral item.

Although difficult to pin down, the pragmatic function of *quello* can be inferred from its association with co-referent topical or focal elements, as in examples (6), (7), (8), (9) and (10), (11) respectively. Moreover, it can refer to a situation in the common ground when occurring in athetic sentence (cf. (12) and Fig. 11). In such pragmatically marked contexts, *quello* is also pitch accented and followed by a prosodic boundary, suggesting that it may maintain the prosodic features of the discourse-salient pronoun along reanalysis. The proposed origin of the pragmatic use of the pronoun *chillo/quello* rather than its use as demonstrative determiner is also in line with the wider tendency in Romance of third person pronouns developing into pragmatic expletives (Corr, 2017; Hinzelin, 2009). Parallel examples can be found in Portuguese (*ele*, Carrilho, 2008), Dominican Spanish (*ello*, Hinzelin and Kaiser, 2007) and Balearic Catalan (*ell*, Bartra-Kaufmann, 2011; Remberger and Hinzelin, 2009). On the other hand, *quello/chillo* as a demonstrative determiner does not show the same ambiguity, nor is pitch accented or phrased apart from the co-referent DP.

The development of *chillo* (< Lat. ECCU ILLU(M)), which should form the basis for the pragmatic use of *quello* examined in this article, can thus be schematized as follows:

- (17) demonstrative determiner > subject pronoun > pragmatic left-peripheral pronoun.

7. Conclusions

The results presented in this article show that social media are a valuable source of spoken data, which can offer new insights into linguistic analysis despite the limitations presented by the low acoustic quality of the signal. By using *Pyrlato*, a novel pipeline for the extraction of spoken data from the internet (cf. Magistro, Crocco, 2023), we collected a corpus of INI featuring real-world speech from talk-shows, movies, theatre, interviews etc. This data collection allowed us to explore the prosodic properties of pragmatic *quello*, a regional feature typical of informal dialogical speech which seems otherwise difficult to elicit through widely used paradigms for prosodic and syntactic research (task-oriented dialogues, Discourse Completion Task, questionnaires, interviews, etc.).

Through auditory and visual inspection of the scraped data, it was shown that *quello* is frequently marked with a high/rising pitch accent when used as a pronoun or as a pragmatic marker. The data support the hypothesis put forward in previous

studies concerning the different prosodic facets characterizing demonstrative determiner *quello* and pragmatic *quello*, in terms of prominence and phrasing. This has led us to formulate a new hypothesis about the origin of the pragmatic uses of *quello/chillo*, namely that it derives from the use of *chillo* as a pronoun. Although real-world data scraped from the internet have a high potential for prosodic and linguistic research, they also have limitations. These have emerged in the case of RQ2. In fact, social media data did not help in casting light on finer differences possibly differentiating between *quello* co-referent with a focal and *quello* co-referent with a topical DP. This limitation highlights the importance of data triangulation in prosodic research (Vanrell, Feldhausen, & Lluïsa Astruc, 2018).

Of course, the assumed parallel between dialectal prosody and the prosody of *quello* in INI deserves a more in-depth investigation. This matter will be addressed in future work on *chillo/quello* specifically (cf. Groothuis, Crocco, 2023, in prep.), as well as more generally for Campania (Leo in prep.). However, we believe that the results of the prosodic inspection of *quello* reported in this article provide us with valuable first insights into the prosodic properties and the historical developments of pragmatic pronouns in varieties spoken in Campania.

Acknowledgements

The authors wish to thank the audience of the AISV conference in Lecce for helpful discussions on the topic, as well as two anonymous reviewers for their valuable comments and remarks. All remaining errors are the authors' own responsibility. This work was financially supported by the Research Foundation – Flanders (FWO), project numbers 12C0723N (Kim A. Groothuis), G010122N (Claudia Crocco & Giovanni Leo), and G004519N (Anne Breitbarth, Claudia Crocco & Giuseppe Magistro).

References

- BARONE, M. (in prep). *Intonation in Pescara: The description of two languages in contact*. PhD dissertation, Università degli studi di Lecce/Universitat Pompeu Fabra.
- BARTRA-KAUFMANN, A. (2011). Recycled neuter expletive pronouns and the nature of the left periphery: *Ell* and relatives revisited. In *Catalan Journal of Linguistics*, 10, 185–219. <https://doi.org/10.5565/rev/catjl.37>
- BECKMAN, M.E. (1997). A typology of spontaneous speech. In SAGISAKA, Y., CAMPBELL, N. & N. HIGUCHI (a cura di), *Computing prosody: Computational models for processing spontaneous speech*. New York: Springer, 7–26.
- BERRUTO, G. (1987 [2012]). *Sociolinguistica dell'italiano contemporaneo* (2nd ed.). Rome: Carocci.
- BOERSMA, P. (2001). *Praat*, a system for doing phonetics by computer. In *Glott. Int.*, 5(9), 341–345.

- BOERSMA, P., Weenink, D. (2021). PRAAT: doing phonetics by computer [Computer program]. Version 6.1.40. <https://www.praat.org/>
- CARRILHO, E. (2008). Beyond doubling: Overt expletives in European Portuguese dialects. In Barbiers, S., Koenenman, O., Lekakou, M. & Ham, van der, M. (a cura di) *Microvariation in syntactic doubling*. Bingley: Emerald group, 301–323. https://doi.org/10.1163/9781848550216_012
- CATALDO, V. (2023). *The intonation of closely related varieties. The case of Campania Italian*. PhD thesis, Università degli Studi di Salerno/Ghent University.
- CHAFE, W.L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View in Subject and Topic. In Li, C.N. (a cura di), *Subject and topics*. New York: Academic Press, 25–55.
- CHOMSKY, N. (1981). *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.
- CORR, A. (2017). The grammaticalization of epistemicity in Ibero-Romance: Alike processes, unlike outcomes. In *Journal of Historical Linguistics*, 7(1), 48–76. <https://doi.org/10.1075/jhl.7.1-2.03cor>
- CROCCO, C., GILI FIVELA, B., & D'IMPERIO, M. (2022). *Comparing prosody of Italian varieties and dialects: Data from Neapolitan*. In FROTA, S., CRUZ, M., & VIGÁRIO, M. (a cura di), *Speech Prosody 2022*, Lisbon, 23-26 May 2022, 140–144. <https://doi.org/10.21437/SpeechProsody.2022-29>
- DE BLASI, N., FANCIULLO, F. (2002). La Campania. In CORTELAZZO, M., MARCATO, C. (a cura di), *I dialetti italiani. Storia, struttura, uso*. Turin: UTET, 628–678.
- D'IMPERIO, M. (2001). Focus and tonal structure in Neapolitan Italian. In *Speech Communication*, 33(4), 339–356. [https://doi.org/10.1016/S0167-6393\(00\)00064-9](https://doi.org/10.1016/S0167-6393(00)00064-9)
- D'IMPERIO, M. (2002). Italian intonation: An overview and some questions. In *Probus*, 14(1). <https://doi.org/10.1515/prbs.2002.005>
- FUCHS, R., MAXWELL, O. (2016). The Effects of mp3 Compression on Acoustic Measurements of Fundamental Frequency and Pitch Range. In BARNES, J., BRUGOS, A., SHATTUCK-HUFNAGEL, S. & VEILLEUX, N. (a cura di.) *Speech Prosody 2016*, 523–527. ISCA. <https://doi.org/10.21437/SpeechProsody.2016-107>
- GAETA, L. (2014). Soggetto o topic? Chillo 'o fatto è nire, nire. In CUGNO, F., MANTOVANI, L., RIVOIRA, M. & SPECCHIA, M.S. (a cura di), *Studi linguisticiti in onore di Lorenzo Massobrio*. Turin: Istituto dell'Atlante Linguistico Italiano, 465–479.
- GROOTHUIS, K.A. (2022). Reconsidering the pragmatic use of chillo in Campania. Paper presented at *Going Romance 2022*, Universitat Autònoma Barcelona.
- GROOTHUIS, K.A., CROCCO, C. (2023). The prosodic features of pragmatic *chillo/quello* in Neapolitan and Campanian Italian: implications for syntax and diachrony. Paper presented at the 17th Cambridge Italian Dialects Syntax Morphology Meeting, University of Zürich.
- GROOTHUIS, K.A., CROCCO, C. (in prep.). The prosodic features of pragmatic *chillo/quello* in Neapolitan and Campanian Italian: implications for syntax and diachrony. Ms, Ghent University.

GROOTHUIS, K.A., CROCCO, C., DI SALVO, M. & MATRISCIANO-MAYERHOFER, S. (2024). Pragmatic “quello” from dialect to regional standard. Paper presented at the Romance Grammars, context and contact 2024, University of Birmingham.

HANSON, H.M. (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. In *The Journal of the Acoustical Society of America*, 125(1), 425–441. <https://doi.org/10.1121/1.3021306>

HINZELIN, M.-O. (2009). Neuter pronouns in Ibero-Romance: Discourse reference, expletives and beyond. In KAISER, G.A., REMBERGER, E.-M. (a cura di), *Proceedings of the Workshop «Null-subjects, expletives, and locatives in Romance»*. Konstanz: Universität Konstanz, 1–25.

HINZELIN, M.-O., KAISER, G.A. (2007). El pronombre ello en el léxico del español dominicano. In MIHATSCH, W., SOKOL, M. (a cura di), *Language contact and language change in the Caribbean and beyond*. Frankfurt am Main: Lang, 171–188.

HORN, L. (1984). Towards a new taxonomy for pragmatic inference: Q- and R-based implicature. In SCHRIFFIN, D. (a cura di), *Meaning, form, and use in context: Linguistic applications*. Washington, DC: Georgetown University Press, 11–42.

KIRBY, J.P., LADD, D.R. (2016). Effects of obstruent voicing on vowel F0: Evidence from “true voicing” languages. In *The Journal of the Acoustical Society of America*, 140(4), 2400–2411. <https://doi.org/10.1121/1.4962445>

LEDGEWAY, A.N. (2000). *A comparative syntax of the dialects of Southern Italy: A minimalist approach*. Oxford: Wiley-Blackwell.

LEDGEWAY, A.N. (2009). *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.

LEDGEWAY, A.N. (2010). Subject licensing in CP: the Neapolitan double-subject Construction. In BENINCÀ, P., MUNARO, N. (a cura di), *Mapping the left periphery*. Oxford: Oxford University Press, 257–296. <https://doi.org/10.1093/acprof:oso/9780199740376.003.0009>

LEDGEWAY, A.N. (2013). Testing linguistic theory and variation to their limits: The case of Romance. In *Corpus*, 12, 271–327. <https://doi.org/10.4000/corpus.2408>

LEO, G. (in prep.). Intonation and rhythm of dialects and Italian. The case of Campania. PhD thesis, Ghent University.

MAGISTRO, G., CROCCO, C., (2023). Pyrlato: a novel methodology to collect real-world acoustic data. In *Estudios de Fonética Experimental*, 35 (ahead of print).

MATURI, P. (2002). *Dialetti e substandardizzazione nel Sannio Beneventano*. Frankfurt: Lang.

MATURI, P. (2023). *Napoli e la Campania*. Bologna: Il mulino.

O’ROURKE, E. (2005). *Intonation and language contact: A case study of two varieties of Peruvian Spanish*. PhD dissertation, University of Illinois at Urbana-Champaign.

RADTKE, E. (1997). *I dialetti della Campania, edizione italiana a cura dell’autore, con la collaborazione di Paolo Di Giovine e Franco Fanciullo*. Roma: Il Calamo.

REMBERGER, E.-M., HINZELIN, M.-O. (2009). *Expletive vs. Discourse marker: Sardinian ello and Balearic Catalan ell*. Talk given at the 38th Romance Linguistics Seminar, Trinity Hall, Cambridge. Trinity Hall, Cambridge.

- RIZZI, L. (1997). The fine structure of the left periphery. In HAEGEMAN, L.M.V. (a cura di), *Elements of grammar*. Dordrecht: Kluwer, 281–337. https://doi.org/10.1007/978-94-011-5420-8_7
- RIZZI, L., BOCCI, G. (2017). Left periphery of the clause. Primarily illustrated for Italian. In EVERAERT, M.B.H., RIEMSDIJK, van, H. (a cura di), *The Wiley Blackwell Companion to Syntax* (2nd ed.). Oxford: John Wiley & Sons, 1–30. <https://doi.org/10.1002/9781118358733.wbsyncom104>
- SORNICOLA, R. (1996). Alcune strutture con pronomi espletivi nei dialetti italiani meridionali. BENINCÀ, P., CINQUE, G., DE MAURO, T., & VINCENT, N. (a cura di), *Italiano e dialetti nel tempo. Saggi di grammatica per Giulio C. Lepschy*. Rome: Bulzoni editore, 323–340.
- TELMON, T. (1994). Gli italiani regionali contemporanei. In SERIANNI, L., TRIFONE, P. (a cura di), *Storia della lingua italiana: Vol. 3: Le altre lingue*. Turin: Einaudi, 597–626.
- TUCKER, B.V., MUKAI, Y. (2023). *Spontaneous Speech* (1st ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108943024>
- VANRELL, M.D.M., FELDHAUSEN, I., & LLUISA ASTRUC. (2018). The Discourse Completion Task in Romance prosody research: Status quo and outlook. In FELDHAUSEN, I., REICH, U., VANRELL, M.D.M. (a cura di), *Methods in prosody: a Romance language perspective*. Berlin: Language Science Press. <https://doi.org/10.5281/ZENODO.1441345>
- VITOLO, G. (2006). L'uso del dimostrativo espletivo in contesto di frasi principali e dipendenti nella parlata salernitana. *Quaderni del Dipartimento di Linguistica - Università di Firenze*, 16, 61–77.
- WARNER, N. (2012). Methods for studying spontaneous speech. In COHN, A., FOUGERON, C. & HUFFMAN, M. (a cura di), *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press, 621–633.