

---

# WHAT WE LEARN ABOUT LANGUAGE FROM SPOKEN CORPUS LINGUISTICS?

## QUÈ PODEM APRENDRE DE LA LLENGUA AMB LA LINGÜÍSTICA DE CORPUS ORALS?

---

MIRIAM VOGHERA  
*Università di Salerno*  
voghera@unisa.it

**Abstract:** Over the last few decades, the Spoken Corpus Linguistics (SCL) has achieved a great deal in terms of quantity and quality of works (O’Keeffe & McCarthy 2010). Enormous progress has been made in the last thirty years and the increment of multimodal corpora stimulates sophisticated investigations on the relationship between the verbal and non-verbal component of spoken communication (Knight 2011). The SCL is a very vital field of research, which is able to provide essential data and tools for the advancement of language knowledge. In this article I will focus on the contribution that SCL and the resulting data provide to general linguistics. In § 2, I discuss the contribution that the SCL gives to a better understanding of linguistic variation; in § 3, I show how the SCL can improve the descriptive adequacy of grammar; finally, § 4 is dedicated to the contribution that speech data can give to a better knowledge of the grammaticality of languages. Across the article I will use mainly data from Italian corpora, but widely validated by comparison with data from corpora of other languages.

**Key words:** Spoken Corpus Linguistics, Speech Mode, Grammar, Italian.

**Resum:** Al llarg de les darreres dècades, la Lingüística de Corpus Orals (LCO) ha avançat força en termes de quantitat i qualitat d’estudis (O’Keeffe & McCarthy 2010). En els darrers trenta anys s’ha assolit un gran progrés com mostra l’increment de corpus multimodals existents i les diverses investigacions sofisticades que aquests corpus han estimulat sobre la relació entre el component verbal i no-verbal de la comunicació oral (Knight 2011). La LCO és un camp de recerca molt vital que pot proporcionar dades i eines essencials per a avançar en el coneixement lingüístic. En aquest article m’ocupe de la contribució que la LCO i els resultats que ha produït proporcionen a la lingüística general. En l’apartat § 2 analitze la contribució de la LCO a una millor comprensió

de la variació lingüística; en § 3 mostre com la LCO pot millorar la descripció gramatical; i, per acabar, l'apartat § 4 s'ocupa de la contribució que les dades orals poden fer a la comprensió de la gramaticalitat de les llengües. Al llarg de l'article utilitze dades que provenen fonamentalment de corpus italians, però també d'altres llengües; aquesta comparació interlingüística permet validar àmpliament les dades.

**Key words:** Lingüística de Corpus Orals, modalitat parlada, gramàtica, italià.



## 1. INTRODUCTION

Over the last few decades, the Spoken Corpus Linguistics (SCL) has achieved a great deal in terms of quantity of work, described languages and quality of tools produced (O'Keeffe & McCarthy 2010). Although speech corpora are still much less numerous than written corpora,<sup>1</sup> there is no doubt that enormous progress has been made in the last thirty years not only in text collection, but also in systems of storage and text-to-speech alignment, which allow multi-level queries. In addition, there is an increment of multimodal corpora that stimulates sophisticated investigations on the relationship between the verbal and non-verbal component of spoken communication (Knight 2011). The SCL is, in short, a very fecond field of research, which is able to provide essential data and tools for the advancement of language knowledge. In fact, in this article I will focus on the contribution that SCL and the resulting data provide to general linguistics. In paragraph 2, I discuss the contribution that the SCL gives to a better understanding of linguistic variation; in paragraph 3, I show how the SCL can improve the descriptive adequacy of grammar; finally, paragraph 4 is dedicated to the contribution that speech data can give to a better knowledge of the grammaticality of languages. Across the article I will use mainly data from Italian corpora, but widely validated by comparison with data from corpora of other languages.<sup>2</sup>

1. It is worth considering that in the British National Corpus, as well as in the Corpus de Referencia del Español Actual, two of the most important projects in the construction of national reference corpora, the proportion between spoken and written texts is one to ten.

2. For a detailed comparison between the corpora of Italian and those of other languages, see Voghera 2017.

## 2. SPOKEN CORPUS LINGUISTICS AND VARIATION

There is no immediate and direct link between sociolinguistics and SCL (Mair 2009; Baker 2010). However, the design of any corpus is the mirror of our hypotheses of linguistic potential variations along various dimensions: diastratic, diaphasic, diatopic, etc. Speech studies have introduced modal dimension, of which the relationship with the more traditional sociolinguistic categories was initially little known on a large scale. This new point of view has required new analyses and a general reconsideration of many parameters, which brought to a better knowledge of all the factors of variation (Biber 1988). The case of Italian is of particular interest because the SCL has been deeply intertwined with sociolinguistic investigations and has allowed, on the one hand, a better description of the sociolinguistic space of Italian and, on the other, a better definition of the notion of spoken language (*parlato*) as such.<sup>3</sup>

Since the 16th century, a condition of diglossia has prevailed in Italy as a result of the presence of «High» and «Low» varieties, corresponding respectively to literary Italian, used almost exclusively in writing, and Italo-Romance dialects, used in speech<sup>4</sup> (Lepschy *et al.* 1996). Therefore, for centuries, speech and writing have differentiated themselves not so much as different modes of communication, but because they have been associated with different languages (De Mauro 2011). It was therefore inevitable that the first investigations into speech would be interwoven with the description of the different diatopic varieties of Italian and with their dialectal substratum (Sornicola 1981; Sabatini 1985; Berruto 2012). The distinction between speech and writing has thus partially overlapped with that between standard Italian and non-standard varieties of Italian: the first associated with writing, the second with spoken communication. This has not allowed to draw a clear boundary between linguistic strategies belonging to the spoken language and structures that are frequent in spoken texts, but are not exclusive to them. In particular, constructions have been attributed to speech that seems to actually belong to social (*italiano popolare* 'popular Italian') or situational varieties more than to different modes of communication. Although these varieties are actually used in speech rather than in writing, they do not reflect the multiplicity of spoken uses, which manifest characteristics that do not belong to a single variety of Italian. For example, many researches have used corpora of popular or socially 'Low'

3. The Italian term *parlato* covers the meanings of the English expressions speech and spoken language.

4. It is worth recalling that what are called dialects do not derive from Italian, nor are they varieties or adaptations of the national language. On the contrary, Italian, along with all the other Italo-Romance dialects, such as Turinese, Milanese, Neapolitan, etc., are all «sister» languages derived from spoken Latin.

connotated Italian, favouring the identification of speech with low, or marginal, registers of the language. If we scroll through, for example, the list of corpora used by Koch & Oesterreicher (1990: 3642) for a comparison of French, Italian and Spanish, out of the ten corpora of Italian consulted by the authors are texts of popular Italian. The term *speech* sometime overlapped with the popular or informal usage, missing the differentiation among the spoken texts and their peculiar features compared to texts of other modes: first of all the written one.

Thanks to the publication of the first national corpus of spoken Italian, it has been possible to acquire more accurate data on the speech of Italians and on the complicated interplay among diatopic, diaphasic and diastratic variables. In 1993 the corpus LIP (De Mauro *et al.* 1993) was published, and today a second updated edition of it is available, the VoLIP (Voghera *et al.* 2014). The corpus consists of about 500,000 word tokens for 60 hours of recordings, collected in four different Italian cities (Milan, Florence, Rome, and Naples) and in five different diaphasic situations (face-to-face conversations; telephone conversations; interviews, debates, and classroom interactions; lectures, sermons, and public speeches; radio and television shows).<sup>5</sup> While the number of samples is variable, the corpus presents a balanced total number of words per city as well as per diaphasic situation. The corpus is freely available on the portal <[www.parlaritaliano.it](http://www.parlaritaliano.it)> and gives parallel access to the acoustic and textual material. The corpus is searchable for a) textual and register variables, as annotated in a metadata IMDI format (<[www.mpi.nl/imdi](http://www.mpi.nl/imdi)>) and b) lexical and morpho-syntactic criteria. The two kinds of queries can be cross-checked and produce orthographic transcriptions aligned with audio files. Lexical and morpho-syntactic queries result in all the texts presenting the requested item (word form, lexeme, part of speech) which is provided with the frequency of occurrence per city and per register. A word form can be singled out in the text and listened to. Word form sequences are also searchable, for example, «ho dormito» («I have slept»), «è una città che» («is a town that»), and the output can be both the transcription or the audio.

The analysis of VoLIP made it possible to distinguish the sociolinguistic correlates of speech from the functional ones. The former consist of constructions that pertain to speech because of their social yield, i.e. because they belong to varieties suitable for the situation in which the speakers communicate; the latter consist of constructions and strategies that, regardless of social and geographical factors, are

5. Other spoken corpora are now available, for instance KIParla (<<https://kiparla.it>>), but the VoLIP remains the one which presents the widest geographical variation in different speech situations; for more details see <[www.parlaritaliano.it](http://www.parlaritaliano.it)> and Cresti & Panunzi (2013).

used because they allow a better working-out of the production and elaboration of texts. From a sociolinguistic point of view, speech mainly correlates, obviously, with the use of the native language, that is, with a variety of immediate and easy social access for the speakers, that guarantees the maximum communicative proximity or narrowness between them. In Italy this variety coincides, depending on the situation, with dialect, with the regional varieties of Italian, with informal registers of standard Italian or with mixed texts (Berruto 2012). In VoLIP, the largest number of dialect forms are to be found in Naples, followed by Rome, Florence and Milan. As expected, the minimum use of dialectal expressions is in public speeches, but the maximum is in telephone conversations.

These findings are interesting because telephone calls allow a degree of intimacy that encourage the local and family linguistic usage. There is no doubt that face-to-face free communicative exchanges between two or more speakers are not only quantitatively prevalent, but also linguistically primary: conversation represents the most natural type of text among the spoken usage of language (De Mauro 1971). It can be considered the primary modelling system underlying any linguistic use: the deep semiotic structure of our linguistic competence. However, it should be stressed that naturalness does not necessarily correspond to informality. While the first refers to the relationship between linguistic structures and enunciative constraints, as well as cultural ones, the second refers to the degree of social distance/narrowness and role between speakers in a given communicative situation: a text can occupy a high point on the scale of naturalness, but it can also be very formal, and vice versa. In our societies, non-primary and less natural spoken uses have developed and consolidated, and they have acquired an important role in the linguistic community which contribute to forming the spoken usage as a whole and also influence the language of conversation. This is confirmed by the linguistic behaviour in telephone calls that take place in an unnatural situation, since the speakers do not have visual control, but can be very informal.<sup>6</sup> Furthermore, the multiplicity of communicative situations

6. The complex interplay between the naturalness of the communicative situation and the levels of formality can be seen also in numerous other types of text, for example, the electoral speeches or in the speech of radio and television talk shows. In the latter the speakers must keep under control a double recipient: the interlocutor in the studio and the audience at home. This produces communicative behaviours different from those of normal conversation, since the speakers in general regulate their own utterances not only on the reactions of the present recipient, but also on the presumed reactions of a vast and largely anonymous public. These are therefore conversations, in which the naturalness of the exchange is often constructed more than spontaneous: it is also presumed that these are texts partially planned in advance on the basis of notes. This does not exclude, however, that the language can also be very informal.

often gives rise to complex texts in which, for example, the dialogue is interrupted, leaving room for long monologic parts, just as situations of unidirectional exchange can include dialogical interludes.

In order to draw a picture of the overall dynamics of all the variables at stake, it is therefore desirable to consider all the forms of spontaneous speech, even the less primary and less natural ones. In my opinion, this is a necessary condition if one wants to distinguish the sociolinguistic correlates of speech specific to a community<sup>7</sup> from the functional ones, highly shared interlinguistically.

## 2. SPOKEN CORPUS LINGUISTICS AND GRAMMAR

From the very early days, SCL has highlighted the distance between spoken usage and grammatical descriptions, based essentially on written usage in formal prose.<sup>8</sup> Today, data on an increasing number of languages show that the functional correlates of speech consist of a coherent and systematic framework of linguistic constructions widely shared at the interlinguistic level, (Biber 1995; Briz 1998; Miller & Weinert 1998; Biber *et al.* 1999; De Mauro 1994; Albano Leoni & Giordano 2005; Cresti & Moneglia 2005, Voghera 1992, 2017). They do not consist of special grammatical paradigms, different rules of word formation or eccentric syntactic forms; rather, they entail specific choices within the language of grammar, which are more adequate to the mode of communication.

Not all linguistic choices are equally probable in every mode of communication because mode is a complex mechanism in which several factors work together: the channel, the type of interaction between speakers and the processes of production and reception. Each of these factors correlates with the others, that is, variation in each of these factors causes a variation in the others as well. The following figure proposes a schematic representation of the setting of each mode component in the speech mode: the bi-directionality of the arrows indicates that all the elements of the scheme condition each other and there is no superordinate element with respect to the others.

7. Stylistic correlates exist too, that is features that are culturally associated to specific genres of speech and writing, although I will not deal with them in this article.

8. It is important to underline that the traditional structure of grammar also neglects all written uses that do not coincide with those found in formal monologic texts that, roughly, coincide with literary or informative prose. Written texts produced with new media, such as e-mail messages, SMS messages, etc., are quite a different matter and require a broader reflection on the new modes: see Crystal 2001; Tavosanis 2011.

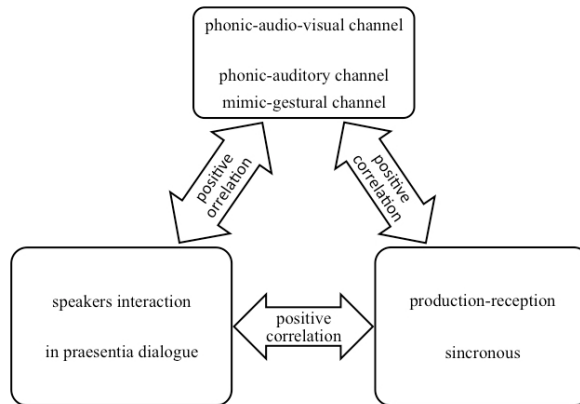


Figure 1: Features of spoken mode of communication

In the speech mode, communication passes through both the phonic-auditory and the mimic-gestural channels. Although both channels can be used separately, in natural communication the two are in such an integrated semiotic relationship that it is preferable to speak of an phonic-audio-visual channel to indicate the normal transmission in human communication (Voghera 2017). There is no spontaneous spoken communication that is not matched by facial expressions, hand movements, postures and body movements, to which I will refer here simply as gestures. We have co-verbal gestures with different functions: to complete the utterance, to substitute part of the utterance, to contradict the verbal sequence, for example in the case of ironic utterances (Kendon 2004; McNeill 2000, 2008; Poggi 2007; Campisi 2018). When we think about the relationship between verbal sequence and gestures we should not imagine that the latter have a merely subordinate function to the word, but rather that there is a relationship of semiotic cooperation between them. The presence of gestures is useful both to the recipient and to the speaker to maintain the speech rhythm and to mark the progression of the information. In fact, gestures are present also when we speak on the phone (Bavelas *et al.* 2008) and among blind speakers from birth, who gesture like and as much as the sighted speakers both when they speak to sighted people and when they speak to each other (Iverson & Goldin-Meadow 2001).

As many works have shown, interpersonal relationships are part of the modal configuration (De Mauro 1982; Halliday 1994; Caffi 2007; Kress 2009; Voghera 2017). The structure of interpersonal relations is the space of intermediation between the properties of the channels and those of the systems of production and perception. In fact, the organization of the relations between producer and recipient strongly







in identifying the relevant points of the dialogue exchange: the Complex Transition Relevance Places.

[...] the fact that there are intonational, syntactic, and pragmatic points of completion means that people can recognize when they are in the midst of a unit that has not yet come to completion in one or more of these ways; in fact, as we have shown, in the majority of cases, not being completed in one of these respects entails not being completed in the other respects either. At a more profound level, our findings point strongly to the need for a thorough linguistic investigation of the ways that intonation, syntax, and conversational action are all *structured* to permit projection of the future course of the unit in progress, including when it can be expected to come to completion. (Ford & Thompson 1996: 171)

All the elements of the speech mode system therefore contribute to conditioning the linguistic choices. Once again, it must be stressed that these are not compulsory choices: there is no automatic and deterministic relationship between the mode of communication and the linguistic product, but rather preferential linguistic choices dictated by the communicative opportunity. These functional linguistic correlates are manifest in speech in natural and spontaneous conditions: it is obvious that in the case of a speech based on a previously prepared written text or of professional speakers the linguistic products can be very different.

There is no level of spoken production that is not conditioned by mode and in fact the SCL has identified textual, syntactic, semantic, lexical, pragmatic functional correlates and so on (Voghera 2017). Although it is not possible here to go into detail and give a description of all the constructions that correlate with speech, it is possible to indicate some general characteristics that they have in common.

The constructions that most frequently positively correlate with the speech mode arise mainly from the need to maintain a good balance between planning/production and reception/processing times. Speakers must save time and energy, without compromising communication success and this happens mainly through a general hypo-specification of the relevant traits for the identification of the elements of the code. In a hypothetical scale that goes from maximum to minimum realization, the distinctive features of the constituents, the spontaneous spoken texts, not only dialogic, tend towards the minimum potential specification at all levels of the structure: phonic, semantic, syntactic, textual (De Mauro 1971). The hypo-specification is balanced by the possibility of making an optimal integrated usage of all available semiotic resources: maximum integration between verbal component and gestures, optimization of the relationship between segmental sequence and suprasegmental level and between verbal sequence and situational context.

Hypo-specification can be observed at phonetic level: spectroacoustic analyses of spontaneous speech clearly show that phonetic realizations do not only present large phenomena of co-articulation, but a general hypo-specification of articulatory targets, up to the contraction or absence of entire phonic sequences, if compared with a hypothetical graphic realization (Albano Leoni & Maturi 1992). Linguistic perception is not in fact a process of recognition, but rather a process of reconstruction and interpretation (Lindblom 2004; Holt & Lotto 2010). In this process, the prosody plays an important role, not only because it does not seem to manifest phenomena of hypo-specification, but also because it seems to play a 'protective' function. In fact, it has been registered that elements in prosodically salient positions, such as those preceding pitch accents or focus accents, are less subject to phenomena of hypo-articulation (Savy 2001).

As one knows, prosody is a powerful device for the production of meaning, which is not isomorphic to other levels of encoding, such as syntax or distribution of information. However, a study made on the Italian spoken corpus Clips (<www.unina.clips>) made it possible to verify the relationship between the phrasing at various levels (Savy & Voghera 2010).<sup>11</sup> We analysed separately the prosodic phrasing of the syntactic units and the syntactic phrasing and constituency of the tonal units and then we compared the two analyses. The prosodic analysis, carried out on an acoustic and auditory basis, consisted of: 1. the segmentation into tonal units; 2. the assignment of a tone on a phonetic basis; 3. the labelling of the pitch accents for each tonal unit. Following a road widely accepted in SCL, we considered the clause the domain of maximum syntactic projecting (Halliday 1985; Chafe 1988; Haiman & Thompson 1988; Miller & Weinert 1998; Biber *et al.* 1999; Voghera *et al.* 2004). The cases of misalignment between the two types of phrasing are very limited, whether we assume the syntactic or the prosodic point of view: only 7.5% of the clauses do not have the left or right boundary that coincides with that of a tonal unit, and only 4.5% of the tonal units do not have the boundary that coincides with a phrase or clause. An important role is naturally played by the distribution of information and the phrases that constitute a tonal unit are often easily identifiable as Topic or Comment. However, there is no isomorphism between syntactic and prosodic levels or between prosodic and informative levels, as shown by the fact that there are tonal unit boundaries that mark disfluencies, reprogramming factors or discursive progression. Also in this case

11. Here I dwell only on the overall data relating to the degree of correspondence of the two types of analysis, while I refer to Savy & Voghera (2010) for details of both analysis.

we can say that prosody has the function of rebalancing because it tends to delimitate meaningful portions on a syntactic or informative level or to isolate sequences that the listener must ignore, such as false starts, disfluencies and all phenomena of reprogramming. What emerges, then, is that a prosodic boundary rarely breaks in a disruptive way a syntactic constituent, phrase or clause, or a constituent that is salient from an informative point of view. The boundaries of the constituents of the various levels tend to be aligned because this eases the production and processing processes, even if there is not a deterministic relationship among the levels.

Another manifestation of hypo-specification, recorded in all SCL studies, is the use of structures with a low intension and a wide extension, which, therefore, create a wide spectrum of meanings and functions, and may occur in a high number of contexts. If there are two competing structures in a given context, one of which may only occur in that context and the other which may occur in more contexts, speech will, in principle, use the second. This depends on the need of planning and speaking at the same time, which reduce the time to look for the most elaborate words or structures: words and multi-functional and polysemic constructions are therefore widely used.

The preference for structures with a wide distribution can be seen at all levels. At the lexical level, synonyms with greater semantic coverage: *macchina* (lit., a generic 'machine') instead of *automobile* ('car'), *arrivare* ('to arrive') instead of *giungere* ('to alight', a more formal expression than 'to arrive') (De Mauro *et al.* 1993). The same principle also applies to the grammatical vocabulary: one prefers to use *ma* (but) instead of *tuttavia* ('however') or *cioè* ('that is') instead of *ossia* ('in other words', more formal than *cioè*), *anche se* ('even if') instead of *quantunque* ('although'), and so on (Voghera 2017). Generally, there is a preference in all spoken texts for the use of connectives that can perform subtle different functions. If we look in particular at the subordinates, we discover that the most frequent types of explicit subordinates have two properties in common: the property of modifying morphologically and functionally different constituents and the amplitude (or vagueness) of semantic values. These two properties are well illustrated both by the relative pronouns and by the conjunction *che*, which are the most frequent subordinators in spoken texts (Voghera 1992). The relative pronouns can relativize in fact any constituent of a sentence represented by a nominal element - subject, object, prepositional phrases - and in Italian allow the use of indicative, conjunctive, conditional and infinitive. The conjunction *che* shows an even greater syntactic flexibility; in fact, it can modify not only nominal phrases, but also verbal phrases, adverbs, entire clauses and even interjections. This corresponds to a zero lexical value that allows this subordinator to cover a wide spectrum of meanings.

It is to these properties that we owe the label of universal unmarked subordinator, assigned to it by Lehmann (1988).

Finally, also in the choice of verbal tenses and modes we tend to use as much as possible those which can assume multiple function as, for instance, in Italian, the indicative present: in VoLIP texts it covers 53% of the verbal forms of the first 2000 lemmas in order of frequency, which in turn cover about 91% of the whole corpus.

Polysemy and low specification at the semantic level are counterbalanced at the textual level with an increase in redundancy, repetition and use of discourse markers. Discourse markers can, in fact, be used for the construction and interpretation of the various portions of the text or to signal the attitude of the speaker towards the message or the interlocutor (Ajmer & Vanderberger 2011; Maschler & Schffrin 2015). In natural spontaneous speech a text which is coherent from the beginning to the end is rarely produced; much more often, the speaker must re-motivate and give coherence to what s/he says during the process. In some cases, therefore, discourse markers are used as a framing device to delimit portions of speech that are textually and pragmatically significant. Other discourse markers' uses occur in dialogues in order to avoid (too many) disfluencies, with an inevitable fragmentation effect and the consequent need for repair mechanisms. They can in fact fulfill the function of covering programming times when there is an imbalance between them and those of verbal production. In this way, the verbal sequence is not properly altered and there is no need for reconstruction on the part of the receiver.

The mechanisms described above are exemplified by the comparison between two narratives, one oral and one written, of the same speaker of a scene in C. Chaplin's film *The Kid*.<sup>12</sup>

12. The example is taken from the corpus Ferrante (2016) collected for the project Modokit (Voghera *et al.* 2020), which compares oral and written narratives of the same speakers in order to build tools for observation and didactic intervention on the relationship between modes of communication and language choices.

I have used here a broad prosodic transcription: // indicates the boundary of the tone units, ↗ indicates a continuation tune with a slightly rising tone, ↓ indicates a falling tone, ↑ indicates a rising tone, ↓↑ a falling-rising tone, → indicates a level tone. Other symbols are: sp=short pause; lp= long pause; vl= vowel lengthening.

Spoken Narrative	Written Narrative
<p>allora↓//quindi&lt;sp&gt; il film che ho visto^//  è un film tratto da&lt;a&gt; &lt;ehm&gt;^// da Charlie  Chaplin^// intitolato il Monello del millenove-  centoventuno &lt;sp&gt;^// ed è&lt;èè&gt; &lt;ehm&gt; diciamo  s'intitola appunto Il Monello sottratto a Charlot  ↓//&lt;sp&gt; e parla di questo bamb+ //cioè almeno  la&lt;aa&gt; la scena iniziale ↓↑// &lt;sp&gt; con cui si  apre il film^// &lt;ehm&gt; mostra chiaramente che  &lt;ehm&gt; come questo bambino venga sottratto^//  diciamo&lt;oo&gt; rapito^// in qualche modo sottrat-  to a Charlot↓// &lt;sp&gt; e &lt;ehm&gt; diciamo che la  scena&lt;aa&gt;^// perlomeno la&lt;aa&gt; la scena iniziale  del film è ambientata in una stanza ↓//&lt;sp&gt; in  una camera ↓↑// &lt;sp&gt; e all'improvviso fanno  irruzione in questa camera proprio dei&lt;ii&gt;^//  insomma non so delle persone non so chi siano...</p> <p><b>English translation</b></p> <p>well then &lt;sp&gt; so the film I saw is a film taken  from &lt;ehm&gt; by Charlie Chaplin entitled the  Kid from nineteen-twenty-one &lt;sp&gt; and is&lt;vl&gt;  &lt;ehm&gt; let's say it is entitled precisely The Kid  taken away from Charlot &lt;sp&gt; and speaks of this  child + that is at least the&lt;vl&gt; the initial scene  &lt;sp&gt; with which the film opens &lt;ehm&gt; shows  clearly that &lt;ehm&gt; how this child is taken away  let's say&lt; vl &gt; kidnapped in some way removed  from Charlot &lt;sp&gt; and &lt;ehm&gt; let's say that the  scene&lt; vl &gt; at least the&lt;vl&gt; the initial scene of  the film is set in a room &lt;sp&gt; in a room &lt;sp&gt;  and suddenly break into this room just some&lt;  vl &gt; well I don't know people I don't know who  they are...</p>	<p>Il film appena visto si intitola «Il Monello» (1921)  di Charlie Chaplin. La scena con cui si apre il  film riguarda questo monello che viene sottratto a  Charlot dalla polizia. I poliziotti fanno irruzione  in questa camera nella quale i due si trovano.</p> <p><b>English translation</b></p> <p>The film just seen is titled «The Kid» (1921) by  Charlie Chaplin. The scene with which the film  opens concerns this kid who is taken away from  Charlot by the police. The cops break into this  room where the two of them are.</p>

The differences between the two texts are very numerous, and the spoken text even appears agrammatical in absence of listening, that is when the verbal sequence is not shaped into a prosodic form.

Looking at the features discussed so far, we notice that the spoken narrative is much longer and redundant; the speaker typically proceeds in a spiral way through progressions, partial backward returns and reformulations:

- (i) mostra chiaramente che <ehm> come questo bambino venga sottratto diciamo<oo> rapito in qualche modo sottratto a Charlot.  
'(it) clearly shows that <ehm> how this child is taken let's say<vl> kidnapped in some way taken from Charlot'

There are many instances of uncertainty expressed by the phenomena of disfluency, parentheses and personal comments, many discourse markers function as hedges to limit the scope or the illocutionary force of the utterances: *diciamo* ('let's say'), *insomma* ('well'),<sup>13</sup> *non so* ('I do not know'). On the whole, the information of the spoken narrative is tortuous and diluted if compared to the compactness and brevity of the written one. In short, the two texts are an emblematic example of the grammatical intricacy and the high lexical density and compactness, indicated by Halliday (1985) as a typical feature respectively of spoken and written language.

These characteristics of spoken texts derive from the dynamic construction of the verbal sequence, which develops in an additive and incremental way (Thompson, Fox & Couper-Kuhlen 2015). The strong temporal constraints imposed on both production and speech processing does not allow for long sentences with numerous levels of embedding. In these conditions, the relationship between linearity, strongly favoured in a situation governed by time, and the syntactic relations, which can normally be discontinuous and can involve several hierarchical levels, must be optimised. For this reason, short constituents are chosen, which follow one another additionally, and hierarchical relationships based on contiguity are favoured. In these cases, the connection can be marked by segmental material, but more often also by the prosodic form. As already said, in the absence of listening the syntactic constituents may appear simply as a list, but they are always prosodically connected or separated. Prosody can play a syntactic role through variations of the melodic extension within the tonal units and those related to specific forms of the intonation profile, which can signal different types of syntactic relations, such as coordination, dependence or juxtaposition (Giordano & Voghera 2009). Moreover, it has been demonstrated that the use of pitch accents and boundary tones, generally defined as continuation tunes, can be functional to the syntactic connection in many languages (Giordano 2008). These generally have a rising tone configuration that is associated with a stressed syllable within the tone unit in the case of accents, at the right end of the tone unit in the case of the boundary. These intonation forms are so frequent in languages that they are one of the topics of discussion on the universality of prosody. Therefore, prosody balances, also at the syntactic level, the possible segmental hypo-specification.

<sup>13</sup>. *Insomma* means literally 'in conclusion', but a more adequate translation in this context is *well* with a hedge function.

It is well known that in speech there are many verbless clauses. SCL studies record for both English and the main Romance languages (Italian, French, Spanish, Portuguese) an average of about 38% of verbless clauses, a percentage that reaches 45% in informal dialogical speech (Biber *et al.* 1999; Cresti & Moneglia eds. 2005).<sup>14</sup> The phenomenon is not marginal and, above all, not limited to sub-standard registers. Traditionally, syntax recognized the classical predicative verbless clauses, typically dirhematic with or without adjuncts, in which the predicative element and the subject of the predication can be easily identified, such as in the following examples:<sup>15</sup>

- (2) It. *Bella questa!* (lit. ‘Beautiful this.FEM’)

Verbless sequences that cannot be interpreted on the basis of the traditional SN+SV sentence pattern are usually considered as fragments or, at most, as elliptical (Elugardo & Stainton 2005). Actually, there is a large quantity of verbless clauses that are not elliptical, but nevertheless fully syntactically autonomous, such as:

- a) syntactically and semantically isolated phrases that often introduce a new Topic or summarise the subject just developed (Sammarco in press):

- (3) It. *Sindrome spalla mano* (lit. ‘shoulder hand syndrome’)

- b) discourse markers, simple yes/no answers, interrupted strings, exclamations:

- (4) It. *Ecco!* (lit. ‘here!’)

- c) argumental and non argumental-clauses, both embedded in another clause or part of a predication which is distributed over several turns, belonging to either one or more speakers:

- (5) It. *E e credo **che di lì*** (lit. ‘and and (I) believe that from there’)

14. Verbless clauses are not exclusive to the spoken texts, as evidenced by the fact that several studies have dealt with non-verbal sentences in literary and non-literary written texts of different languages (Cresti 2005; Lefeuvre 1999; Ferrari 2002; Progovac 2006). We do not have such extensive data on Italian written texts, however, from estimations on samples of small extension, but very diversified for registers, in the non-dialogic writing we can say that they are around 10% (Voghera & Turco 2008).

15. The predicative verbless clauses are not necessarily composed by a NP subject and a predicative phrase, but can be composed by a NP+ PP or NP + Adv.



Interestingly, all verbless clauses behave like the verbal ones with regard to intonational phrasing, melodic profiles and the selection of pitch accents and tone boundaries to express their syntactic role and their relationships with other clauses. Besides, not only they do not present a reduced prosodic form, but their prosody fully expresses both the illocutive force and the modal values. This confirms the fact that prosody is part of the clause form. This is now confirmed by the studies on implicit prosody, which show how silent prosody plays a crucial role also in processing the comprehension of written sentences (Frazier & Gibson 2015).

In an analysis on a corpus of Italian, English and Spanish of both task-oriented dialogue and spontaneous speech, we collected the following quantitative data on verbless clauses (Landolfi, Sammarco & Voghera 2010).<sup>16</sup>

Table 1: Percentage of verbless clauses in a corpus of task oriented dialogues and spontaneous speech (Landolfi, Sammarco & Voghera 2010)

<b>Dialogues</b>	Italian	English	Spanish
<b>Verbless Clauses</b>	28%	24%	33%

The overall qualitative and quantitative data are consistent with those of other research on the subject. In all three languages, the verbless clauses are about 30% out of the total of the clauses; the most numerous type is that constituted by DM and simple yes/no answers and, as already seen in other corpus-based studies (Cresti 2005), classical predicative clauses represent only 10% out of the total. Elliptical structures, in which it is possible to immediately recover the verb from the previous context, are 12%.

Verbless clauses are particularly well-suited to the semiotic and cognitive conditions in which speech naturally takes place, i.e. rapid and on-line linguistic production/reception. This shows that speakers, according to different communicative situations and exigencies, produce a variety of structures that cannot be easily subsumed under canonical sentencehood representation (De Mauro 1974; Voghera 1992, 2017; Progovac 2006; Sammarco 2017; Heine 2001). This is supported by the higher number of verbless clauses in task-oriented dialogues in all three languages, in which the speakers are engaged in continuous exchanges of questions and answers, due to the specific task.

16. The syntactic annotation was carried out by AN.ANA.S. (*Annotazione e Analisi Sintattica*), a system freely downloadable at the portal <www.parlaritaliano.it>, developed within a treebank project for the syntactic annotation of both spoken and written texts (Voghera & Cutugno 2014). The system is constituent-based and allows the organization of syntactic units within a hierarchical structure, according to XML standards. Constituent relations are coded directly using elements nesting XML properties.

Speech syntax therefore has different forms of clauses by constituency, with a high percentage of light clauses, i.e. consisting of a single phrase. This is reflected in the distribution of the parts of speech. Several SLC of different languages found that in spoken texts there is a greater frequency of verb tokens. While the frequency of verbs positively correlates with the frequency of adverbs and conjunctions, the frequency of nouns correlates with that one of adjectives and prepositions. These correlations illustrate the role of nouns and verbs as attractors of modifications and potential nuclei of syntagmatic units of various extension and nature (phrase- clause).

These trends are clearly evident when one compares the distribution of the tokens for each part of speech in the VoLIP with that one in two corpora of written Italian (Voghera 2017): one consisting of novels (PTLI, De Mauro 2007) and one consisting of newspapers (CS).<sup>17</sup> The noun area goes from 36% of occurrences in speech to 42% in novels and reaches 56% in newspapers. The difference between the two corpora of writing is to be attributed to the fact that the novels include many varied texts, even with dialogic parts. On the contrary, the CS corpus represents the monological expositive writing *par excellence*, in which nominalized syntax has its highest expression, of which I give here an example:

- (6) Il ministro dell'Interno ha provveduto a stabilire le linee guida ed il formulario per la presentazione delle domande di contributo, i criteri per la ripartizione e per la verifica della corretta gestione del medesimo contributo e le modalità per la sua eventuale revoca.

'The Minister of the Interior has established the guidelines and the form for the submission of applications for contributions, the criteria for the distribution and verification of the correct management of the same contribution and the procedures for its possible revocation.'

Extraordinarily similar data are found in the British National Corpus (BNC) (Rayson, Wilson & Leech 2001; Leech, Rayson & Wilson 2014).<sup>18</sup> Also in the BNC, there is a strong correlation between the use of nouns and prepositions, especially in informative writing, and verbs and conjunctions, especially in dialogic speech. In this case, an internal variation of speech and writing emerges, strongly connected to the presence/absence of dialogue and to the purpose of the text, in which fiction is halfway between conversational speech and expository prose.

17. The CS corpus was not published, but data and information can be found in De Mauro *et al.* 1993.

18. In Leech *et al.* (2014) data on parts of speech are derived from a subsection of the British National Corpus consisting of one million of written words and one million spoken words.

Thus, the variable that most strongly affects the different frequency of the parts of speech is certainly the presence of speech turns, i.e. the opposition dialogue vs. non-dialogue. A first element that can determine the low frequency of nouns in dialogue is the sharing of the context and of the universe of reference by the speakers (cf. Lambrecht 1994). Normally, in dialogues, there is a continuous reference to contextual elements, much more frequently than in other types of text. The reasons are well known: the interlocutors find it much more natural and easy to use the elements of the context rather than words. We know, in fact, that under equal conditions, brevity or economy are rewarded in dialogical communication. Besides nouns seem to entail a higher cognitive weight in programming process, as shown by a significant lower rate of speech in their production in spontaneous speech. A very recent study on nine different languages shows that in all nine languages there is a significant different speech rate in the production of nouns and verbs: the articulation rate before nouns is around 3.5% slower than before verbs and the probability of pauses before nouns is about 60% greater than before verbs (Seifart *et al.* 2018). This involves replacing nouns with deictic and anaphoric elements such as personal and possessive pronouns or demonstrative pronouns. After all, pronouns in speech are more than double that in novels and four times more numerous than in newspapers. The use of deictic elements instead of nouns is, moreover, somehow obligatory in spontaneous dialogues to the point that if the deictic and/or anaphoric elements were replaced by nouns, we would obtain unacceptably redundant texts. Since one of the main semantic-discursive functions of nouns is to introduce the referents of discourse and to recall them, it is natural that dialogue has few of them. The frequency of nouns seems to increase in parallel with the need to create a universe of reference shared among the interlocutors, that is, when the need to use terms with full reference increases.

Another factor that conditions the frequency of nouns is the management of information distribution. The use of N is in fact often correlated to the expression of the Topic (Ferrari *et al.* 2008: 84). Of course, what differentiates the dialogue from other situations is once again the possibility of introducing a Topic even in the absence of a full referential expression. This is what is commonly called contextual deixis: the possibility of introducing and/or topicalizing elements that are recognizable to the speakers for their salience in the given communicative situation (Berretta 1994; Vanelli & Renzi 1995). Here is a conversation of VoLIP, which exemplifies very well the possibility of topicalizing contextually given and salient elements without ever naming them:

## (7) VoLIP-MA2I

1. A: //chiedo scusa ma sotto c'era così tanta gente→//
2. B: //prego↓//
3. A: //per per cancellazione di iscrizione mi han detto che bisogna fare certe cose↓//
4. B: //sì↓//
5. A: //ma la domanda va fatta su carta libera <sp> normale dichiaro di non aver e↗//  
ah non ho non ho messo su gli altri occhiali↓// dunque marca da bollo quattromila lire↓// ah **questa qui sarebbe**↗//
6. B: //sì **questa qua è la domanda**↓// deve farla vistare di là all'ufficio prestiti↗//  
[...]
7. B: //dichiaro↗// vede↑//mhm mhm↓//
8. A: //dichiaro e gli faccio compilare **questo**↓//
9. B: //poi de<ve>↗// intanto si faccia firmare **questo che è qua**↓//
10. A: // già che son qui mi faccio firmare **questo**↓// e poi dopo quando è tutto↗//
11. B: //consegna giù in segreteria↓//

1. A: I beg you pardon but downstairs there were so many people
2. B: please
3. A: to to reopen for cancellation of enrollment they told me that certain things must be done
4. B: yes
5. A: but the application must be made on normal unstamped paper I declare that I have not # ah I have not I have not put my glasses therefore a four thousand liras official stamp ah this one here it would be
6. B: yes but this is the application it has to be approved by the Loan Office from the other side  
[...]
7. B: I declare you see mhm mhm
8. A: I declare and have him fill out this form
9. B: then you must in the meantime get this one signed which is here
10. A: since I'm here I'll get this one signed and then later when it's all over
11. B: you bring it down to the office

This is a conversation between a lady and an employee in a university administrative office. Apparently there are two documents that the lady must prepare: the application ('la domanda') and a form, which the son must sign. The first reference to the application is in the end of turn 5 with a demonstrative pronouns *questa sarebbe* ('it would be this.FEM one here') to which corresponds a full reference in the next turn *questa è la domanda* ('this.FEM is the application'). The form, instead, is never

mentioned and is always referred to by deictic pronouns *questo* ('this.MAS'): lines 8, 9, 10. Since the object is physically present during the conversation, there is no need to use a noun and this is not an obstacle for becoming the Topic of the following turns.<sup>19</sup>

In summary, the distribution of the parts of speech responds to the needs of semantic, pragmatic and syntactic economy. In speech, where the texts are mainly made up of turns of dialogue, whose duration is necessarily very variable and unpredictable, the speakers must optimise the verbal resources and exploit as much as possible the non-verbal semiosis or, better, the natural interaction between modal conditions and verbal construction. Monologic spoken texts, on the contrary, are usually more planned and are addressed to recipients who do not have the same right to take the ground: it follows, of course, that in a monologue we can grant a greater explicitness, which is also achieved through the greater use of nouns.

Once again, the speech mode encourages speakers to make hypo-specific choices. As I hope to have shown with the functional correlates presented in this paragraph, this choice does not affect the periphery of the linguistic system, but all levels of linguistic organization. However, it is not common to find systematic description of these phenomena in grammars, even scientific ones, unless they are specifically dedicated to the comparison between speech and writing (Biber *et al.* 1999).

Prosody is completely ignored outside specialist studies, just as the pragmatic dimension is neglected. Yet, it became clear that these are two levels of pervasive significance in communication. When we analyse spoken texts without prosody, they are not only incomprehensible, but nearly ungrammatical. The role of the producer and of the recipient couple is equally decisive, as they must collaborate not only for the transmission of the propositional content, but also for the construction of the interaction, which constitutes the necessary cement for any communication. In analysing a spoken text, it is evident that the transferred content conveys constant traces of the subjectivity of the producer and of the continuous effort to enter into a relationship with the recipient. Moreover, in speech, the interpersonal relationship is in many cases the objective of the communicative exchange. In short, the construction of a good intersubjective relationship is often at the same time the medium and the aim of our discourses. The very interplay between the propositional and pragmatic levels is not, of course, exclusive to the spoken mode and it would be a mistake to believe that the pragmatic dimension distinguishes speech from the other modes of communication. However, it takes on unavoidable importance in dialogue because of the physical closeness of the participants in the communication and the continuous exchange of roles.

19. For the sake of brevity, I do not consider other indexical phenomena present in this conversation.

In conclusion, the speakers have very regular and interlinguistically shared behaviours. The inclusion of these uses in a general grammar of a language is necessary not only to account for the spoken communication, but to improve the adequacy of the description: it would make it possible to capture new phenomena and to refine the definition of phenomena already known both on a synchronic and on a diachronic level. It is well known that for some years now diachronic investigations have been open to the study of non-continuous texts in order to better comprehend the incidence of pragmatic factors in linguistic change (Traugott 2003). Moreover, the quantitative incidence of some structures in speech has stimulated the comparison with different diachronic phases: in Italian some diachronic investigations have clearly shown that many linguistic structures considered typical of spoken texts were actually already attested in ancient Italian (D'Achille 1990).

### 3. SCL AND LANGUAGE GRAMMATICALITY

The analysis, even if synthetic, of some sociolinguistic and functional correlates of the speech mode highlights some characteristics of languages that should be taken into account not only in the study of spoken texts. The choice of a mode deeply affects the composition and reception of texts, and switching from one mode to another does not involve a simple change in the outside shell, but a process that touches all levels of text. These changes do not concern the use of a different grammar, but rather a reformulation of the message. Every intermodal passage consists, rather than in a change of grammar, in a real re-semiotisation (Kress 2009). All three components of the mode (channel, interpersonal relation, production and elaboration processes), which correlate with each other, contribute to these results.

This vision entails that the modal components are not external elements to the signification. This undermines the idea, common to many grammatical models, of a pre-packaged language code that is completely external at the time of enunciation to the relationship between users. The most common idea in grammars is that there is a canonical conformation of signs, represented by their complete formal and semantic form. This canonical conformation would disclose in the most redundant way the identifying features of the signs, and, as such, manifests all their necessary categorical properties, which are identifiable on the basis of a closed, synchronically defined and stable number of attributes. The comprehension between speakers should take place through the recognition of these attributes with a relatively low cost procedure, as for example happens with the units of a computation. Even when features are not

considered inherent, it is believed that they are always retrievable within a given context. It is evident that the implicit model of reference is formal writing, in which the phenomena of fuzziness and gradience are marginal and superficial.<sup>20</sup> But it is also evident that the functional correlates of speech undermine this vision of the code and of the communicative process.

Firstly, the internal data of speech and the comparison between speech and writing clearly show that each linguistic element can be realized at different levels of specification: phones, parts of speech, clauses can be expressed with all possible categorical features or only with a minimum portion of them. Speakers, rather than relying on very pre-packaged signs, find it more convenient to evaluate each time the weight of the different variables and - through a calculation of probabilities and a game of specification and continuous refinement - come to encode and decode the messages (Lindblom 2004; Piantadosi, Tily & Gibson 2011). In fact, the different degrees of specification of a sign do not derive from and do not cause a random and chaotic variability, but are conditioned by precise choices of communicative efficiency. Thanks to SCL, we now have the possibility of measuring this variability and having consistent data on many different languages, on the variation ranges and on the correlation between the variations of the different levels. This variance is a stable and necessary datum for communication as whole and, therefore, not confined to low registers. The most realistic representation of the code units is consequently not that of elements with inherent and stable properties, but of elements with flexible and intrinsically gradient boundaries. This is attested, among other things, by the role of the prosody, gradient by its nature, which is one of the most significant modal correlates because it is co-essential to any textual choice. This is obviously true in spoken texts, but there is now evidence that prosody is projected onto written texts even in silent reading and influences the parsing and the interpretation of a written sentence: what is called implicit prosody (Fodor 2002; Breen 2015).

Secondly, speech data, phonetic and linguistic perception research all testify that not only can constituents be hypo-specified, but that this is the favorite sign condition in natural and spontaneous communication. All functional correlates of speech tend towards hypo-specification, except prosody: although it is possible to vary the level of specification in the realization of the elements of the code, speakers choose the lowest one. This is possible and convenient precisely because the speakers are not users outside the code, but act as code makers, as members of the language community and as

---

20. Informal handwriting of course presents many phenomena of hypo-specification similar to those observed in speech; for a survey of non-categoricity and gradience see Aarts *et al.* 2004.



authors of single communicative acts. The enunciative phase is not a mere execution of a complete product, but it is the beginning of a cognitive, linguistic and social process that has at its center the couple producer / listener. Regardless of how much the speakers are aware of it, the pragmatic dimension is not accessory nor external to the code: in other words, there is no linguistic production and comprehension that can take place without an active and constructive relationship between speakers. As soon as we move away from fictitious grammar examples and deal with corpus data, not only spoken, language becomes a field of open negotiation between speakers. A grammaticality of languages emerges that includes as its constituent element the pragmatic dimension, which should not be considered as external to the code, but as an integral part of its significant potential. As real data analyses increase it is becoming more and more clear that pragmatic values are internal to many categories and lexical uses values: I have already mentioned the role that the consideration of pragmatic meanings has gained in diachronic studies, for example in the study of grammaticalisation processes (Davidse *et al.* 2010).

Elements that are central for the functioning of the linguistic system emerge: they are difficult to observe in other modes, but nevertheless are crucial for the general grammar of a language. Think, for example, of the relations between prosody, syntax, information structure and pragmatics or between segmental phonetic realization and linguistic perception. It is clear that this type of observation cannot only be described, but also requires a reconsideration of the grammaticality of languages as systems deeply linked to the intersubjective relationship between the speakers that makes possible a flexible and elastic use of the relevant features of linguistic signs. This, we have seen, allows for very different levels of specification, connected to the choice of mode and of course to the sociolinguistic variables: from the minimum of conversation between family members speaking the same sociolinguistic variety to the maximum of specification required in formal scientific prose. Neither of the two productions is more canonical than the other and neither of the two is more grammatical than the other. The assumption of this point of view allows us to construct a general grammar that is able to explain speech within a unitary framework, and not as an eccentric and/or marked appendix. Although the interpenetration between modal and code elements takes place in every linguistic activity, speech offers the advantage of observing the process as it unfolds, without any subsequent normalization, to which linguistic uses are often subjected in other modes of communication.

Summarising, the data on spoken languages, acquired thanks to the SCL, allow a more realistic view not only of the *parole*, but also of the linguistic system in itself and its potentiality: they enrich the knowledge of linguistic variation as well as the

flexibility of the system, of its components or levels. Besides, the SCL offers methods of measuring the frequency and distribution of the variation and its correlation with the different dimensions of speaking. This interpenetration between qualitative and quantitative analyses provides the necessary basis for a more adequate explanatory models of grammar and makes SCL's studies an invaluable resource to successfully accomplish this objective.

MIRIAM VOGHERA  
*Università di Salerno*  
 voghera@unisa.it

ORCID 0000-0002-5121-6819

## SYMBOLS AND ABBREVIATIONS

lp = long pause	→ = level tone
sp = short pause	↗ = continuation tone with a slightly rising tone
v1 = vowel lengthening	↓ = falling tone
// = boundary of tone unit	↑ = rising tone
↓↑ = falling-rising tone	

## BIBLIOGRAPHIC REFERENCES

- AARTS, B. *et al.*, eds. (2004) *Fuzzy grammar: A reader*, Oxford, Oxford University Press.
- AIJMER K. & A. M. VANDENBERGEN (2011) «Pragmatic Markers», in J. Zienkowski, J.-O. Östman, J. Verschueren (eds.), *Discursive Pragmatics*, Amsterdam-Philadelphia, John Benjamins, p. 223-47.
- ALBANO LEONI, F. & R. GIORDANO, eds. (2005) *Italiano parlato. Analisi di un dialogo*, Napoli, Liguori.
- ALBANO LEONI, F. & P. MATURI (1992) «Per una verifica pragmatica dei modelli fonologici», in G. Gobber (ed.), *La linguistica pragmatica*, Roma, Bulzoni, p. 39-49.
- BAKER, P. (2010) *Sociolinguistics and corpus linguistics*, Edinburgh, Edinburgh University Press.
- BAVELAS, J. *et al.* (2008) «Gesturing on the Telephone: Independent Effects of Dialogue and Visibility», *Journal of Memory and Language*, 58, p. 495-520.

- BERRETTA, M. (1994) «Il parlato italiano contemporaneo», in L. Serianni, P. Trifone (eds.), *Storia della lingua italiana*, vol. II: *Scritto e parlato*, Torino, Einaudi, p. 239-70.
- BERRUTO, G. (2012 [1987]) *Sociolinguistica dell'italiano contemporaneo*, Roma, Carocci
- BIBER, D. (1988) *Variation across speech and writing*, Cambridge, Cambridge University Press.
- (1995) *Dimensions of Register Variation: A Cross-Linguistic Comparison*, Cambridge, Cambridge University Press.
- BIBER, D., S. JOHANSSON, G. LEECH, S. CONRAD & E. FINEGAN (1999), *Longman Grammar of Spoken and Written English*, Harlow, Essex, Pearson Education.
- BREEN, M. (2015) «Empirical Investigations of Implicit Prosody», in L. Frazier, E. Gibson, (eds.), *Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor*, Cham/Heidelberg/New York/London, Springer, p. 177-192.
- BRIZ, A. (1998) *El español coloquial en la conversación*, Barcelona, Ariel.
- CAFFI, C. (2007) *Mitigation*, Amsterdam-Oxford, Elsevier.
- CAMPISI, E. (2018) *Cos'è la gestualità*, Roma, Carocci.
- CHAFE, W. L. (1988) «Linking Intonation Units in Spoken English», in J. Haiman, S. A. Thompson (eds.), *Clause Combining in Grammar and Discourse*, Amsterdam-Philadelphia, John Benjamins p. 1-27.
- CLANCY, P. M. *et al.* (1996) «The Conversational Use of Reactive Tokens in English, Japanese, and Mandarin», *Journal of Pragmatics*, 26, p. 355-87.
- CRESTI, E. (2005) «Enunciato e frase: teoria e verifiche empiriche», in M. Biffi, O. Calabrese, L. Salibra (eds.), *Italia linguistica: discorsi di scritto e di parlato. Scritti in onore di Giovanni Nencioni*, Siena, Prolagon, p. 249-60.
- CRESTI, E. & M. MONEGLIA, eds. (2005) *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages*, Amsterdam, John Benjamins.
- CRESTI, E. & A. PANUNZI (2013) *Introduzione ai corpora dell'italiano*, Bologna, il Mulino.
- CRYSTAL, D. (2001) *Language and the Internet*, Cambridge, Cambridge University Press.
- DAVIDSE K., L. VANDELANOTTE & H. CUYCKENS, eds. (2010) *Subjectification, Inter-subjectification and Grammaticalization*, Berlin, De Gruyter.
- D'ACHILLE, P. (1990) *Sintassi del parlato e tradizione scritta della lingua italiana. Analisi di testi dalle origini al secolo XVIII*, Roma, Bonacci.
- DE MAURO, T. (1971) «Tra Thamus e Theuth. Uso scritto e parlato dei segni linguistici», in *Senso e significato. Studi di semantica teorica e storica*, Bari, Adriatica Editrice, p. 96-114.
- (1974) «Premesse ad una raccolta di tipi sintattici», in M. Medici, A. Sangregorio (eds.), *Fenomeni morfologici e sintattici dell'italiano contemporaneo*, Roma, Bulzoni, p. 551-74.

- (1982) *Minisemantica dei linguaggi non verbali e delle lingue*, Roma-Bari, Laterza.
- (2007) *Primo tesoro della lingua letteraria italiana del Novecento*, Torino, Utet.
- (2011 [1963]) *Storia linguistica dell'Italia unita*, Roma-Bari, Laterza.
- , ed. (1994) *Come parlano gli italiani*, Firenze, La Nuova Italia.
- DE MAURO, T., F. MANCINI, M. VEDOVELLI & M. VOGHERA (1993) *Lessico di frequenza dell'italiano parlato*, Milano, Etaslibri.
- ELUGARDO, R & R. J. STAINTON, eds. (2005) *Ellipsis and Nonsentential Speech*, Dordrecht, Springer.
- FERRANTE, G. (2016) *Parlato e scritto a confronto. Una nuova metodologia*, Tesi di laurea magistrale, Università degli Studi di Salerno, a.a. 2015-16.
- FERRARI, A. (2002) *Valore intrinseco e funzioni testuali della frase nominale*, in H. Jansen et al. (eds.), *L'infinito e oltre. Omaggio a Gunver Skytte*, Odense, Odense Universitetsforlag, p. 171-90.
- FODOR, J. D. (2002) «Psycholinguistics cannot escape prosody», in *Proceedings of the Speech Prosody 2002 Conference*. [Online: <<http://aune.lpl.univ-aix.fr/sp2002/papers.htm>>.]
- FORD, C. E. & S. A. THOMPSON (1996) «Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for The Management of Turns», in E. A. Schegloff, E. Ochs & S. A. Thompson (eds.), *Interaction and Grammar*, Cambridge, Cambridge University Press, p. 134-84.
- FRAZIER, L. & E. GIBSON, eds. (2015) *Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor*, Dordrecht, Springer.
- GIORDANO, R. (2008) «Continuation tunes in two central varieties of Italian: phonetic patterns and phonological issues», in A. Botinis (ed.), *Second ISCA Workshop on Experimental Linguistics*, Athens, Nation and Kapodistrian, p. 97-100.
- GIORDANO, R. & M. VOGHERA (2009) «Frase senza verbo: il contributo della prosodia», in A. Ferrari (ed.), *Sintassi storica e sincronica dell'italiano*, vol. II, Atti del X Congresso della Società internazionale di linguistica e filologia Italiana, Firenze, Franco Cesati, p. 1005-24.
- HAIMAN, J. & S. A. THOMPSON (1988) «Clause Combining in Grammar and Discourse», Amsterdam-Philadelphia, John Benjamins, p. 1-27.
- HALLIDAY, M. A. K. (1985) *Spoken and Written Language*, Oxford, Oxford University Press.
- (1994) «Language as social semiotic», in *Language and literacy in social practice*, p. 23-43.

- HEINE, L. (2001) «Non-Coordination-Based Ellipsis from a Construction Grammar Perspective: The Case of the Coffee Construction», *Cognitive Linguistics*, 22, p. 55-80.
- HOLT, L. & J. L. LOTTO (2010), «Speech Perception as Categorization», *Attention, Perception, & Psychophysics*, 72, p. 1218-27.
- IVERSON, J. M. & S. GOLDIN-MEADOW (2001) «The resilience of gesture in talk: Gesture in blind speakers and listeners», *Developmental Science*, 4, p. 416-422.
- JEFFERSON, G. (1973) «A Case of Precision Timing in Ordinary Conversation: Overlapped Tag-Positioned Address Terms in Closing Sequences», *Semiotica*, 9, p. 47-96.
- KENDON, A. (1967) «Some Functions of Gaze-Direction in Social Interaction», *Acta Psychologica*, 26, p. 22-63.
- (2004) *Gesture: Visible action as utterance*, Cambridge, Cambridge University Press.
- KNIGHT, D. (2011) «The future of multimodal corpora», *Revista brasileira de linguística aplicada*, 11, p. 391-415.
- KOCH, P. (1988) «Italienisch: Gesprochene und geschriebene Sprache», in G. Holtus, M. Metzeltin, C. Schmitt (eds.), *Lexikon der Romanistischen Linguistik (lrl)*, vol. IV, *Italienisch, Korsisch, Sardisch*, Tübingen, Niemeyer, p. 189-206.
- KOCH, P. & W. OESTERREICHER (1990) *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*, Tübingen, Niemeyer.
- KRESS, G. (2009) *Multimodality: A social semiotic approach to contemporary communication*, Abingdon/New York, Routledge.
- LAMBRECHT, K. (1994) *Information Structure and Sentence Form: A Theory of Topic, Focus, and the Mental Representations of Discourse Referents*, Cambridge, Cambridge University Press.
- LANDOLFI, A., C. SAMMARCO & M. VOGHERA (2010) «Verbless Clauses in Italian, Spanish and English: A Treebank Annotation», in S. Bolasco, I. Chiari, L. Giuliano (eds.), *Statistical Analysis of Textual Data. Proceedings of the 10th International Conference Journées d'Analyse statistique des données Textuelles*, Milano, led, p. 1187-94.
- LEECH, G., P. RAYSON & A. WILSON (2014) *Word Frequencies in Written and Spoken English: Based on the British National Corpus*, London, Routledge.
- LEFEUVRE, F. (1999) *La phrase averbale en français*, Paris, L'Harmattan.
- LEHMANN, C. (1988) «Towards a Typology of Clause Linkage», in J. Haiman, S. A. Thompson (eds.), *Clause Combining in Grammar and Discourse*, Amsterdam-Philadelphia, John Benjamins Publishing, p. 181-225.
- LEPSCHY, A. L., G. C. LEPSCHY & M. VOGHERA (1996) «Linguistic Variety in Italy», in C. Levy, (ed.), *Italian Regionalism*, Oxford, Berg Publishers, p. 69-80.

- LINDBLOM, B. (2004) «The organization of speech movements: Specification of units and modes of control», in *From Sound to Sense: 50+ years of discoveries in speech communication*, Conference Proceedings (CD publication), Cambridge, Research Laboratory of Electronics / Massachusetts Institute of Technology, p. 86-97.
- MAIR, C. (2009) «Corpus linguistics meets sociolinguistics», in *World Englishes Problems, Properties and Prospects*, Amsterdam-Philadelphia, John Benjamins, p. 39-60
- MC NEILL, D. (2008) *Gesture and thought*, Chicago, University of Chicago Press.
- , ed. (2000) *Language and gesture*, Cambridge, Cambridge University Press.
- MILLER, J. E. & R. WEINERT (1998) *Spontaneous spoken language: Syntax and discourse*, Oxford, Oxford University Press.
- O'KEEFE, A. & M. MCCARTHY, eds. (2010) *The Routledge Handbook of Corpus Linguistics*, Abingdon/New York, Routledge.
- PIANTADOSI, S. T., H. TILY & E. GIBSON (2012) «The Communicative Function of Ambiguity in Language», *Cognition*, 122, p. 280-91.
- POGGI, I. (2007) *Mind, hands, face and body. A goal and belief view of multimodal communication*, Berlin, Weidler.
- PROGOVAC, L. (2006) «The Syntax of Nonsententials: Small Clauses and Phrases at the Root», in L. Progovac *et al.* (eds.), *The Syntax of Nonsententials: Multidisciplinary Aspects*, Amsterdam-Philadelphia, John Benjamins, p. 33-71.
- RAYSON, P., A. WILSON & G. LEECH (2001) «Grammatical Word Class Variation within the British National Corpus Sampler», in *Language and Computers*, 36, p. 295-306.
- SABATINI, F. (1985) «L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane», in G. Holtus, G. (ed.) *Gesprochenes Italienisch in Geschichte und Gegenwart*, Tübingen, Gunter Narr Verlag, p. 154-184.
- SACKS, H., E. A. SCHEGLOFF & G. JEFFERSON (1974), «A Simplest Systematics for the Organization of Turn Taking for Conversation», *Language*, 50, p. 696-735.
- SAMMARCO, C. (in press) «Online text mapping: the contribution of verbless constructions in spoken Italian and French», in C. Mauri, E. Gorla, I. Fiorentini (eds.), *Building categories in interaction: linguistic resources at work*, Amsterdam/Philadelphia, John Benjamins.
- SAVY, R. (2001) «L'interfaccia tra livelli di analisi del parlato: rapporti tra riduzioni segmentali e schemi prosodici», in F. Albano Leoni *et al.* (eds.), *Dati empirici e teorie linguistiche*, Bulzoni, Roma, p. 309-28.
- SAVY, R. & M. VOGHERA (2010) «A Corpus-Based Study on Syntactic and Phonetic Prosodic Phrasing Boundaries in Spontaneous Italian Speech», in M. Hasegawa-



- Johnson (ed.), *Proceedings of Speech Prosody*. [Online: <<http://speechprosody2010.illinois.edu/papers/100077.pdf>>.]
- SEIFART, F., J. STRUNK, S. DANIELSEN, I. HARTMANN, B. PAKENDORF, S. WICHMANN, A. WITZLACK-MAKAREVICH, N. H. DE JONG & B. BICKEL (2018) «Nouns slow down speech across structurally and culturally diverse languages», *Proceedings of the National Academy of Sciences*, 115, p. 5720-5725. [Online: <<https://doi.org/10.1073/pnas.1800708115>>.]
- SCHIFFRIN, D. & Y. MASCHLER (2015) «Discourse Markers: Language, Meaning, and Context», in D. Tannen, H. E. Hamilton & D. Schiffrin (eds.), *The handbook of discourse analysis*, Oxford, Blackwell, p. 189-221.
- SORNICOLA, R. (1981) *Sul parlato*, Bologna, il Mulino.
- STIVERS, T. *et al.* (2009) «Universals and Cultural Variation in Turn-Taking in Conversation», in *Proceedings of the National Academy of Sciences of the United States of America*, 106, p. 10587-92.
- TAVOSANIS, M. L. A. (2011) *L'italiano del web*, Roma, Carocci.
- THOMPSON, S. A., B. A. FOX & E. COUPER-KUHLEN (2015) *Grammar in Everyday Talk: Building Responsive Actions*, Cambridge, Cambridge University Press.
- TRAUOGOTT, E. (2003) «From Subjectification to Intersubjectification», in R. Hickey (ed.), *Motives for Language Change*, Cambridge, Cambridge University Press, p. 124-39.
- VANELLI, L. & L. RENZI (1995) «La deissi», in L. Renzi, G. Salvi, A. Cardinaletti (eds.), *Grande grammatica italiana di consultazione*, vol. III, Bologna, il Mulino, p. 261-376.
- VOGHERA, M. (1992), *Sintassi e intonazione nell'italiano parlato*, Bologna, il Mulino.
- (2017) *Dal parlato alla grammatica*, Roma, Carocci.
- VOGHERA, M. & F. CUTUGNO (2010) «ANA.NA.S.: Aligning Text to Temporal Syntagmatic Progression», in M. Mahlberg, V. González-Díaz, C. Smith (eds.), *On-line Proceedings of 5th Corpus Linguistics Conference*. [On line: <http://ucrel.lancs.ac.uk/publications/cl2009/>, art. 209.]
- VOGHERA, M., G. BASILE, D. CERBASI & D. FIORENTINO (2004) «La sintassi della clausola nel dialogo», in Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (eds.), *Il parlato italiano*, Napoli, D'Auria Editore. [CD-ROM, B17.]
- VOGHERA, M., C. IACOBINI, R. SAVY, F. CUTUGNO, A. DE ROSA & I. ALFANO (2014) «VoLIP: a searchable Italian spoken corpus», in L. Veselovská, M. Janebová (eds.), *Complex Visible Out There*, Proceedings of the Olomouc Linguistics Colloquium: Language Use and Linguistic Structure, Olomouc, Palacký University, p. 627-640.



- VOGHERA, M., V. MAYRHOFER, M. RICCI, F. ROSI & C. SAMMARCO (2020) «Il Modokit. Un identikit modale delle produzioni parlate e scritte dalle medie al biennio», in M. Voghera, P. Maturi & F. Rosi (eds.), *Orale e scritto, verbale e non verbale: la multimodalità nell'ora di lezione*, Firenze, Cesati, p. 225-245
- VOGHERA, M. & G. TURCO (2008) «Il peso del parlare e dello scrivere», in M. Petto-  
rino *et al.* (eds.), *La comunicazione parlata*, Atti del Congresso internazionale,  
Napoli, Liguori, p. 727-60.
- VOLIP: <<http://www.parlaritaliano.it/index.php/it/volip>>.