Spoken corpora and pragmatics
Corpora orais e pragmática

Massimo Moneglia*
University of Florence
Firenze / Italy

ABSTRACT: The goal of this paper is to present arguments in favour of two points related to the study of oral corpora and pragmatics: a) at the level of annotation, corpora must ensure the parsing of the speech flow into utterances on the basis of prosodic cues and provide an easy access to the acoustic source; b) at the level of sampling, corpora must ensure the maximum representation of context variation, rather than speaker variation. We will present the reasons which support the very basic prosodic annotation of speech (prosodic boundaries) as a means to obtain relevant data from the speech flow. Starting from our present knowledge about the distribution of speech acts types in spoken corpora, we will present the reasons why building corpora in accordance to a context variation strategy should expand our knowledge of pragmatics. Additionally, we will claim that prosody is the necessary interface between locutive and illocutive acts and we will show that a deeper prosodic analysis is necessary to grasp unknown speech act types from language usage. Finally, we will briefly sketch the main assumptions of the Language into Act Theory (CRESTI, 2000) which is dedicated to the link between prosody and pragmatics and helps make explicit core aspects of pragmatic knowledge.

KEYWORDS: oral corpora; pragmatics; annotation; sampling; speech act types; prosody; Language into Act Theory.

RESUMO: O objetivo deste artigo é apresentar argumentos favoráveis a dois pontos relacionados ao estudo de corpora orais e pragmática: a) no nível da anotação, os corpora devem garantir o processamento do fluxo discursivo em enunciados, baseando-se em chaves prosódicas, e oferecer fácil acesso aos arquivos de som; b) no nível da amostragem, os corpora devem garantir a representatividade máxima de variação contextual, ao invés de variação de falantes. Apresentaremos os motivos que sustentam a escolha das fronteiras prosódicas como o referencial básico para a anotação prosódica da fala, como uma forma relevante de se obterem dados importantes do fluxo discursivo. Partindo do nosso conhecimento atual sobre a distribuição tipológica de atos de fala em corpora orais, apresentaremos as razões pelas quais a construção de corpora de acordo com a estratégia da variação contextual deve expandir o nosso conhecimento sobre pragmática. Adicionalmente, defenderemos que a prosódia é a interface necessária entre atos locutórios e ilocutórios e mostraremos que uma análise prosódica mais profunda é necessária para que se obtenham atos de fala desconhecidos a partir do uso da língua. Por fim, esboçaremos rapidamente os principais pressupostos da Teoria da Língua em Ato (CRESTI, 2000), a qual se debruça sobre a ligação entre a prosódia e a pragmática e auxilia na explicitação dos principais aspectos do conhecimento pragmático.

PALAVRAS-CHAVE: corpora orais; pragmática; anotação; amostragem; tipologia dos atos de fala; prosódia; Teoria da Língua em Ato.

* moneglia@unifi.it

1. Introduction

In my view, in order for spoken corpora to be exploited in a way that will enhance our knowledge in the domain of pragmatics to be built, two main basic strategies should be followed: a) at the level of annotation, corpora must ensure the parsing of the speech flow into utterances on the basis of prosodic cues and provide an easy access to the acoustic source; b) at the level of sampling, corpora must ensure the maximum representation of context variation, rather than speaker variation. These criteria, which have been applied in the construction of the C-ORAL-ROM corpus (CRESTI; MONEGLIA, 2005) and have been in practice at the LABLITA lab at the University of Florence, have ensured a good basis for grounding pragmatic concepts on actual speech data (CRESTI, 2000; CRESTI; FIRENZUOLI, 2001; FIRENZUOLI, 2003; SCARANO, 2003; FROSALI, 2006, CRESTI; MONEGLIA, 2010; CRESTI; MONEGLIA; TUCCI, in press).

The goal of this paper is to present arguments in favour of these two choices. In 2 we will present the reasons which support the very basic prosodic annotation of speech (prosodic boundaries) as a means to obtain relevant data from the speech flow. In 3, starting from our present knowledge about the distribution of speech acts types in spoken corpora, we will present the reasons why building corpora in accordance to a context variation strategy should expend our knowledge of pragmatics. In 4, we will claim that prosody is the necessary interface between locutive and illocutive acts and we will show that a deeper prosodic analysis is necessary to grasp unknown speech act types from language usage. In 5 we will briefly sketch the main assumptions of the Language into Act Theory (CRESTI, 2000) which is dedicated to the link between prosody and pragmatics and helps make explicit core aspects of pragmatic knowledge. According to this theory it is possible to identify the components of the utterance responsible for the illocutionary activity (Comment Unit) and to get clear distinctions between the main pragmatic functions allowed by the language structure, i.e., illocutionary activity and dialogue regulation activity.

More generally, in this paper, we will argue that the possibility to get robust knowledge about language structures that govern speech act performance in the ordinary use of language depends on a better understanding of the link between prosody and pragmatics. This relation and the need for a corpus-based strategy in pragmatic studies are both fundamental steps for grounding pragmatics on strong empirical evidence.

2 . Basic prosodic annotation for the exploitation of spoken corpora

2.1. Pragmatic units of reference for spoken language and prosody

If pragmatics is to profit from the huge amount of evidence which can be derived from contemporary corpus linguistics, these corpora must provide language data which are proper objects for pragmatic analysis; i.e., units of reference within the corpus which show pragmatic qualities. The series of lexical entries which constitute the speech flow (wording) do not provide this minimal linguistic entity directly.

In the case of written language, the nature of the linguistic units ranking above word level is clear. Although it may be chosen at different levels, i.e. argument structures, sentences or clauses, or head dependent structures (ABEILLÉ, 2003), written language can be properly parsed according to syntactic and semantic principles. Conversely, the identification of the units of reference in a spoken corpus can hardly be identified through the same syntactic and semantic devices (BLANCHE-BENVENISTE, 1997; BIBER *et al.*, 1999; CRESTI, 2000; MILLER; WEINERT, 1998; IZRE'EL, 2005).

Reference units for spontaneous speech are commonly identified with the term "utterance". The utterance might be anchored to syntactic and/or semantic properties as well. For instance, it can be identified with a syntactic *clause* (MILLER; WEINERT, 1998), or, as in The Longman Grammar, with a *C-Unit* with or without a clause structure (BIBER, *et al.*, 1999). Clair Benveniste proposed to identify the nucleus of an utterance in a macro-syntactic domain based on a *noyau* bearing a modal value (BLANCHE-BENVENISTE, 1997; BENVENISTE *et al.*, 1990). The definition of such an entity is a complex matter when its annotation in the speech flow is required. The main problem is that in spoken language a lot of configurations that are not clauses may turn out to be utterances in the speech flow. Almost 1/3 of speech events, according to the C-ORAL-ROM for the Romance languages and the Longman Grammar for English, do not have a verb and therefore do not show a clear syntactic structure. (BIBER *et al.*, 1999; MONEGLIA, 2005; MONEGLIA, 2006).

The following example taken from the LABLITA corpus of spoken Italian corresponds to one dialogic turn in which one speaker performs a word sequence. Considering the mere linear word sequence, no configuration pattern can be clearly identified and, from a pragmatic point of view, it is not possible to decide what the pragmatic value of any group of words is.

*SUS: *lei gliene serve una anch'a lei una in più o no no lei ha questa*
     [you need one more too or not no you have this one]

According to pragmatic tradition (AUSTIN, 1962), the utterance is *the minimal linguistic entity such that can be pragmatically interpreted*; i.e. the linguistic entity that is 'concluded' and 'autonomous' from a pragmatic point of view (QUIRK *et al.*, 1985; CRESTI, 2000), but pragmatics can hardly benefit from corpus data if the object that carries pragmatic qualities, i.e. the utterance, is not identified in a corpus. For instance the above sequence cannot be interpreted even if one knows the context of the utterance (the speaker has been asked by a professor to make photocopies of a paper).

In this frame a speech event may also be identified as a dialogue act, and recorded in a dialogue representation scheme. This solution has been clear ever since the origin of corpus linguistic studies (see SINCLAIR; COULTHARD, 1975, and the literature cited below), but the task is hard to be undertaken and the identification of dialogue acts are difficult to be agreed upon, given that speech acts are also quite underdetermined (FAVA, 1995; KEMPSON, 1977).

In any event, however, this task necessarily requires considering the acoustic information, since the evaluation of the prosodic performance is crucial to determine the value of a speech act. Therefore, the access to acoustic information is the basic requirement for whatever exploitation of spoken corpora in the domain of pragmatics.

In the above example, the solution could be: "listen to a speech extract and provide your parsing of the speech flow into utterances". But what determines the parsing of the speech flow once the acoustic information is provided? The operative criteria which lead to the annotation of utterance boundaries in the speech flow must be explicit in order for the obtained data to be reliable and consistent for pragmatic and linguistic studies.

Approaches may diverge on this. My point is that the reference unit for spoken language is not underdetermined if pragmatic and prosodic features of speech are taken into account. Classic studies on prosody have always highlighted the fact that utterances end with a terminal profile (CRYSTAL, 1975; KARCEVSKY, 1931) and this quality is clearly perceived by speakers in conjunction with the assignment of an illocutionary value to a stretch of speech. From this point of view, this simple property can be considered a property equivalent to speech acts, to be used as a heuristic to determine the utterance boundaries in the speech flow: each string ending with a perceptively relevant terminal break is an utterance, in principle matching with a speech act (MONEGLIA, 2005).

According to this method, the speech flow and its transcription can be easily parsed. In the above example, when the grouping of words through intonation is considered , the identification of speech act boundaries is "naturally" guaranteed and it turns out that the above mysterious dialogic turn is made up of four utterances (marked by the terminal signs "?" or "//").

*SUS: *lei / gliene serve una anch'a lei ? una in più / o no ? no // lei ha questa //*
   [you / (do) you need one also for you ? one more / or not ? no // you have this one //]

The criterion for the identification of utterance boundaries is intonation-based. This criterion does not imply the evaluation of the different intonation features and their categorization (evaluation of the movement types, tones, levels, focal points), which is very complex, but is only based on perception: detection of terminal and non-terminal prosodic breaks.[1] These cues are so prominent that they require little training to be recognized. Moreover, the experience of corpus annotation has shown that the perception of terminal breaks is consistent at a cross-linguistic level; English, Dutch, Italian, French, Spanish, European Portuguese, Brazilian Portuguese, Hebrew – all have been the object of this annotation with successful results (IZRE'EL *et al.*, 2005; AMIR *et al.*, 2004; MONEGLIA *et al.*, 2005; MONEGLIA *et al.*, 2010; BUHMANN *et al.*, 2002).

This practice allows for the possibility to get low cost information on speech acts from huge amounts of corpus data. It is reliable from the point of view of the detection of utterances in the speech flow. In this approach, the parsing of the speech flow into discrete speech events is not a function of the recognition of a specific speech act type by the labeler in any annotation schema, since the assignment of utterance boundaries is independently motivated. This property is in some sense quite widely recognized. Some spoken dialogue annotation tasks, for instance, the DRI/DAMSL and HCRC system dialog act codings, work under the same assumption (see CARLETTA *et al.,* 1996; JURAFSKY *et al.,* 1997) i.e. the dialogue act labeling and segmentation of 'utterances' is understood to proceed in tandem.

---

[1] Prosodic breaks must not be mixed up with pauses when looking at utterance boundaries. In around 60% of cases, pauses act as a re-enforcement of terminal prosodic breaks; however also around 40% of non terminal breaks are accompanied by a pause. See Moneglia (2005).

The correspondences between labeled Break Indices (i.e. intonational and intermediate phrases) and the majority of dialogue act boundaries are compatible with results from earlier studies about the relationship between intonational features and discourse (e.g. LEHISTE, 1975; NAKATANI *et al.,* 1995; SWERTS, 1997, SHRIBERG *et al.*, 1998). Dialogue act boundaries usually coincided with intonation boundaries in the MAP TASK corpus with matches of 88% for HCRC moves, and 84% for DAMSL dialogue acts (see below).

However, working within this frame, the prosodic boundaries strategy has not been really exploited having as an end the annotation of dialogue acts. Although the coding scheme for dialogue acts provides a closed list of possible moves, a competent speaker may find it difficult to identify and define the performed act. The replicability of the coding scheme is, as a matter of fact, one of the main problems for the annotation of dialogue acts, even in quite restricted domains.

For this reason, once the utterance limits are identified, the language string corresponding to an utterance is the linguistic entity which is suitable for receiving a certain tag. In other words, the definition of utterance limits is a matter of direct perception, while the assignment of a specific value to a dialogue act is a categorization issue, involving our knowledge of linguistic values. One can count speech acts without a clear agreement on their illocutionary value.

The annotation of utterance boundaries, according to the annotation of prosodic breaks perceived with a terminal value, does not go hand in hand with the ability to assign a specific value to an utterance (categorization task), but rather with the judgment that the utterance is an object of interpretation in the world.

In other words, a competent speaker can agree with the fact that the utterance being regarded can be interpreted, but may diverge, for many reasons, as to the specific value to be assigned to the utterance itself. The capacity to assign the quality of "being interpretable in the world" to a stretch of speech follows from this "illocutionary principle" and is a function of perception that is based on unconscious features.

This idea is not foreseen in the Searlian paradigm (SEARLE, 1983), in which intentional activities, such as language understanding in this case, are up to consciousness. Understanding that a stretch of speech is an object of interpretation in the world is not a function of the conscious assignment of a specific interpretation.

2.2. Speech act performance and syntactic relations

The traditional point of view that the reference unit of spontaneous speech can be detected when the relation among words generates autonomous compositional elements is strongly challenged by the prosodic strategy just described. Most scholars partake the view that prosodic criteria must be considered but always in conjunction with syntactic and semantic evidence (BENVENISTE, 1997, Rhapsodie Project). This is reasonable, but may lead to thorny problems if pragmatics is to be taken seriously.

The following argument shows, in particular, that the parsing of the speech flow into discrete units according to the detection of terminal prosodic breaks is not only a necessary condition, but also a sufficient condition. Speech event boundaries can be identified through prosodic boundaries apart from any other semantic or syntactic consideration, since prosodic boundaries are a function of speech act performance.

The underdeterminacy of syntactic structures in spontaneous speech is not only linked to the absence of verbs, but also to pragmatic activity itself. For instance, when a verbal proposition may, in principle, be figured out from the speech data, this not always provides the actual structure of a speech act. The dialogic turns reported below from the French and Portuguese C-ORAL-ROM collection are presented in a bare transcription without prosodic tagging. The two strings show the same superficial structure, which is a verbal nucleus followed by an adverbial expression:

*EMA: ça c' est clair de plus en plus   [this is clear more and more]
*NOE: estive no Chiado há pouco tempo   [I have been in the Chiado recently]

In both cases, on the basis of semantic and syntactic considerations, it is possible to figure out one proposition; i.e. one verbal nucleus with an adverbial extension. However, if the prosodic information provided by terminal breaks is considered, the two strings show different pragmatic properties since only the second one is accomplished within the boundary of one utterance, while in the first case the adverbial expression performs an independent monorematic utterance, and it is an independent 'adverbial clause', as in the following notation.

* EMA: [1] ça c' est clair // [2] de plus en plus //   [this is clear // more and more //]
*NOE: [1] estive no Chiado / há pouco tempo //   [I have been in the Chiado /
                                                                          recently //]

This has consequences on pragmatic grounds: given that two illocutionary activities are accomplished by the speaker, we cannot figure out only one syntactic structure. The idea (quite *a prioristic* indeed) that speech acts are just a matter of performance and that one single syntactic program is "executed in two utterances" is not consistent with the pragmatic interpretation. If pragmatics regards units of reference corresponding to speech acts and their structure, it cannot be admitted that "one speech act performance" is the performance of two speech acts. What is actually performed by the speaker must be taken seriously by the theory. Therefore, in no circumstance will the adverb modify the verb, since it gives rise to an independent reference unit. The syntax of pragmatic units is not independent.

Of course the reverse is also true. Although in principle an adverbial clause can accomplish one utterance, in no circumstance, in the second turn, does the adverb perform an independent act, since it does not follow a terminal break and does not bear any illocutionary value alone. In summary, the access to the prosodic information determines the structure of the speech flow; it does not read the structure of independently motivated semantic/syntactic entities.

3. Speech act variation and the representation of the language usage

3.1 Corpus-based detection of speech act variability

A corpus-based pragmatics must provide the means to specify what the speech acts actually performed in ordinary conversation are, and what differentiated linguistic properties they show. Searle's taxonomy (SEARLE, 1969) is still probably the most influential speech act classification. The taxonomy was set up at the end of the sixties within a logical paradigm and is based on lexical properties. In his conception, the linguistic counterpart of a speech act, i.e. the utterance, is equivalent to a *performative* predicate applied to a proposition [F(p) = u]. On the basis of a "Principle of Expressibility" a correspondence between *speech act* types and performative verbs is established. Speech acts are defined in accordance, as a set of performative verbs belonging to five classes sharing a set of necessary and sufficient conditions of application (SEARLE, 1969).

However, when carrying out corpus-based experimental research, this point of view turns out to be not adequate to capture real data. The richness of the actions carried out in ordinary conversations is not recorded by the list

of performative verbs, and the *Expressibility Principle* does not provide a valid heuristic to detect actual speech acts, which have, almost always, a *primitive* form in spontaneous speech. More specifically, while importance is given to linguistic actions that never occur in corpora or are rare, several – even very common – *speech acts* are not identified, since they have no equivalent performative (*refusal, deixis, call, instructions*). Even more intriguing: in spontaneous speech, although a performative sentence may, in principle, provide possible paraphrases for an utterance in its primitive form, there is no guarantee that the act actually performed belongs to that type.

The general point is that classical speech act theory lacks giving the appropriate value to prosody, which is the real means used in speech to express speech act types. For this reason, our knowledge about the set of speech act types that are possible in language (and their definition) is still very far from a satisfactory state. We will discuss this in 4. Nevertheless, corpus-based studies solely provide the most promising data which can increase our understanding in this domain.

Let us take a look at some findings in corpus-based speech act detection and classification.

A lot of work was being done at the end of the 90's in the domain of Dialogue Act Modeling for Automatic Recognition of Conversational Speech. In the map tasking coding scheme (ANDERSON *et al.*, 1991), the set of possible dialogue acts (moves in the map task) were investigated and the relevant link between prosodic and discourse structures underlined. Stirling *et al.* (2001) studied these moves and their relation to prosody in accordance with the HCRC map task coding scheme (CARLETTA *et al.*, 1996, 1997) and the 'Switchboard' version of the DRI/DAMSL scheme (JURAFSKY *et al.*, 1997) in detail. The following is the set of acts identified in the richer DAMSL coding scheme.

TABLE 1
DAMSL coding scheme

SWBD-DAMSL codes (based on JURAFSKY *et al.*, 1997)	
Forward-Communicative-Functions:	**Backward-communicative-functions:**
Statement	*Agreement*
statement-non-opinion	accept
statement-opinion	accept-part
Influencing-addressee-future-action	maybe
open option	reject-part
yes-no-question (info-request)	reject
wh-question (info-request)	hold
open question (info request)	*Understanding*
or-question (info request)	signal non-understanding
or-clause after y/n question (info request)	acknowledge
rhetorical question (info request)	backchannel in question form
declarative question (info request)	acknowledge-answer
tag question	repeat phrase
action-directive	collaborative completion
Committing-speaker-future-action	summarize/reformulate
offer	appreciation
commit	sympathy
Other-forward-function	downplayer
conventional opening	correct-misspeaking
conventional closing	*Answer*
explicit performative	yes answers
exclamation	no answers
other forward function	affirmative non-yes answers
thanking	negative non-no answers
you're welcome	other answers
apology	no plus expansion
	yes plus expansion
	statement expanding y/n answer
	expansions of y/n answers
	dispreffered answers

However, map task is spontaneous speech recorded in one quite peculiar situation only. Current trends in corpora which document a huge variety of sociolinguistic and pragmatic domains show that the set of possible speech acts may vary in accordance with the variety of contexts that are sampled in the corpus.

Yuki, Abe and Lin (2005) show that 50 types of functions have been identified by the Usage Based Linguistic Informatics Group (UBLI) in Japanese corpora, and more types have been extracted from the other foreign languages. A reduced table of the more frequent acts, derived from the combination and comparison among the previous ones, has been proposed below.

TABLE 2
40 Functions (YUKI; ABE; LIN, 2005)

Greetings	Thanking	Attracting attention	Introducing oneself	Apologizing
Giving	Saying goodbye	Asking Information (price)	Asking Information (experience)	Telling one's plan
Asking Information (degree)	Asking Information (time)	Asking Information (number)	Saying how and why	Asking skill and ability
Asking Information (existence and place)	Asking Information (attribute)	Saying one's opinion	Saying one's taste (thing)	Saying one's taste (behaviour)
Stating procedure and order	Asking what one is	Saying how one acts under certain circumstance	Comparing (comparative and superlative degree)	Suggesting
Confirming duty/negating	Prohibiting	Instructing	Asking for unacceptable thing	Confirming duty / affirming
Inviting	Advising	Demanding	Stating one's hope	Introducing someone

The analysis carried out during the last decade based on our Italian corpora has led to the identification of a larger set of about 90 *speech act types* in speech (CRESTI; FIRENZUOLI, 2001; FIRENZUOLI, 2003).

TABLE 3
LABLITA Corpus Based Reference Table of Speech Acts Classes and Types

REPRESENTATIVES		DIRECTIVES	Announcing	EXPRESSIVES	Complaint	RITES
Concluding	Confirmation	Distal recall - not visible object	Advising	Exclamation	Imprecation	Thanks
Weak assertion	Approval	Distal recall - visible object	Warning	Expression of contrast	Insinuation	Greetings
Answering	Disapproval	Proximal recall	Suggesting	Expression of obviousness	Derision	Apologies
Commentary	Agreement	Distal deixis	Proposal	Softening	Provocation	Welcome
Strong assertion	Disagreement	Proximal deixis	Recommend	Expression of surprise	Reproaching	Congratulation
Identification		Presenting (object/event)	Invite	Expression of fear	Hint	Wishes
Verification		Introducing (person)	Prompt	Expression of relief	Encouragement	Compliments
Claim		Request of information	Permit	Expression of uncertainly	Assuring	Declarations of legal value
Hypothesis / Supposition		Request of action	Authorize	Expression of doubt	Threatening	*condemnation*
Explanation		Order	Prohibition	Expression of Certainty	Giving up	*condolences*
Inference		Total question	Instruction	Expression of wish		*baptism*
Definition		Partial question		Expression of Disbelief		*promise*
Narration		Alternative question		Expression of Pity		*bet*
Describing		Request of confirmation		Irony		
Quotation		Reported speech		Regret		REFUSALS

490   RBLA, Belo Horizonte, v. 11, n. 2, p. 479-519, 2011

The LABLITA research, based on a corpus of 10h; 9300 utterances, ranging over a large variety of informal situations,[2] also showed that 90% of utterances perform a set of roughly 30 speech act types, which are the more common in everyday conversation. The relative frequency of speech act classes in this corpus is the following:[3]
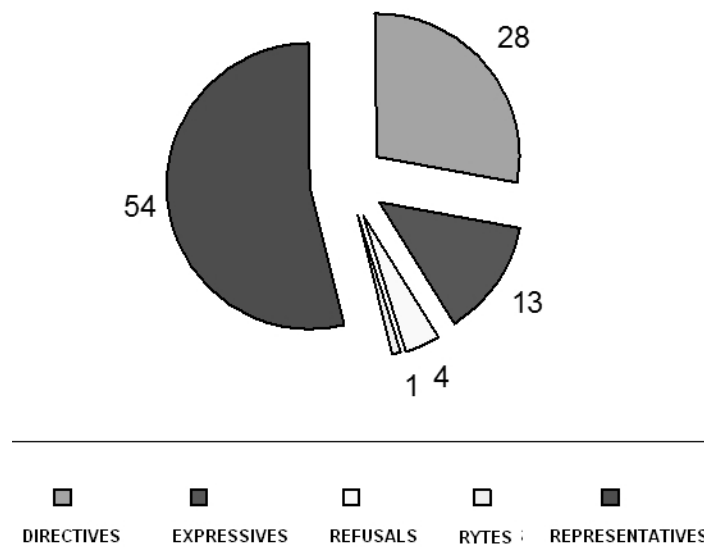

FIGURE 1 - Percentage of utterances of 5 Illocutionary classes in a
sampling of the LABLITA Corpus (FIRENZUOLI, 2003)

Despite the difference in annotation strategy and label choice, these tagsets record at least some similar items. Comparing the three tables we can observe that the LABLITA tagset is larger, but the intersection of recorded speech act types is quite reduced. One easy conclusion: should one wish to represent the spontaneous speech universe in order to capture the variety of possible speech act types, the constitution criteria must ensure the widest

---

[2] Ratio of utterance sampling is 15%.

[3] Despite huge theoretical differences in the definition, the LABLITA illocutionary classes turn out very close to Searle's ones, with minimal adjustment (Commissives are not considered, Declarations are named Rites, an idiosyncratic class "Refusal" which cover a frequent act "NO!" in speech).

possible variation in speech contexts, and the lowest control in the speech event, which is exactly the opposite of what collections that are restricted to a specific task do.

A second requirement follows from the fact that, in the previous LABLITA pie, rites are unduly underrepresented. This is contradicted by the objective frequency of *salutation*, *thanks* and other everyday conventional declarations in daily life. This shortcoming depends on the practical choices made in transcribing samples. The beginning of interactions is almost always avoided since interactions start being more natural (ignoring the recording apparatus) after a while; the end of interactions are almost never sampled, since samples are always shorter than the whole of the interaction. Therefore, if this kind of act should be investigated, the corpus sampling must provide data regarding full pragmatic interactions. This means, more general criteria of corpus design should be integrated with criteria regarding how sessions are sampled. In this case, the map task strategy prevails.

The variety of types of conventional activities allowed by the linguistic system is obviously to be found within the main classes of Representatives, Directives and Expressives which record the highest number of speech act instances and, therefore, contain the relevant variation. This is the main area for future corpus-based research that is, however, strongly dependent on annotation schemas and identification criteria.

3.2 Corpus Sampling

The setting up of spontaneous speech language resources must ensure a huge corpus variety to allow speech act type detection. This requirement is similar to what happens with lexical variation. The representation of a sufficient number of contexts, covering relevant types of speech events in the universe, is the only possible strategy to get data for a frequency lexicon. A high-frequency lexicon may be underrepresented in specific pragmatic domains which, on the other hand, may maximise the probability of occurrence of low-frequency lexical items. The linguistic properties of the speech events vary in connection with non-linguistic variations. The connection between non-linguistic variation and linguistic variation goes beyond the frequency of lemmas: while lexical variation depends on topics, pragmatic variation depends on the needs of the interactive context and on the speaker's personal attitude and habits in that context. The goal to represent the variety of speech acts performed in everyday life from language usage data poses a problem of

representation that is common in corpus linguistics, but is particularly sensitive in the spoken domain. There are relevant technical constraints to speech recording that are not present in written resources. Moreover, speech performance consistently varies from context to context and from individual to individual depending on many parameters.

According to sociolinguistic studies (LABOV, 1966; BERRUTO, 1987; BIBER, 1988; DE MAURO *et al.*, 1993; GADET, 1996) and also to recent initiatives for the annotation of corpus metadata (IMDI), the spontaneous speech universe foresees variation parameters that can be divided into three main groups: a) channel parameters; b) contextual parameters; d) demographic parameters.

Channel variation
   a. Face-to-face interactions in natural contexts
   b. Telephone recordings
   c. New media audiovisual interactions
   d. Human / machine interactions
   e. Media productions
   f. Written to be spoken

Contextual variations parameters
   a. *Structure of the linguistic event*: speech events having a dialogue or a multi-dialogue structure vs. monologues
   b. *Social context*: interactions belonging to *family* and *private life* vs. *interactions taking place in public*
   c. *Domain of use:* domains of social environments, activities and professional domains such as law, business, research, education, politics church, etc.
   d. *Genre:* lesson, debate, chat, row, storytelling, professional explanation, interview *etc*
   e. *Register*: context requirements regarding formal register vs. informal language uses

Demographic parameters: the main sociologic qualities of speakers
   a. Age
   b. Sex
   c. Education

d. Occupation

e. Geographical origin

f. Social class

g. City vs. Country

The impact of such variation parameters on the spontaneous speech universe cannot be pre-theoretically foreseen as for instance in the written part of the BNC. To provide a significant sampling of the population according to demographic parameters and then record them across their lifespan is, in principle, the best strategy. If the socio-demographical sampling of the population is valid, the linguistic sampling will also be valid as far as this population will be recorded through all relevant contexts of the day. All contexts occurring in society will have probability of occurrence according to their frequency in the life of the population and at the same time all language styles and personal variations due to sociologic qualities will be captured.

CoSIH (IZRE'EL *et al.*, 2001) was designed to ensure this. Day-long recordings of 950 informants representing all social and ethnic groups of the Israeli population have been planned over a one-year period. In this procedure informants are captured in recordings while they go through all the contextual and interpersonal situations that occur in the day, so ensuring speech data that are balanced at the same time both at sociological and contextual variation levels.

However, the CoSIH approach is not easily pursued. Indeed, to my knowledge, no corpus has been accomplished at present with this approach. From a practical point of view the recording of most contexts of use requires setting up a recording apparatus beforehand, and those situations remain excluded if not planned. The strategy is also difficult to be applied for legal or moral reasons in countries where the signed agreement of each intervenient in a recording is required beforehand, and the recording of many professional situations like business transactions are constrained by strict rules that go beyond the expressed agreement of the speakers.

If the strategy is not followed coherently the result will have exactly those variations that are significant for speech act variation reduced. For instance the BNC tries to integrate demographic and contextual criteria and dedicates almost half of its spoken part to recordings provided by a significant sampling of the British population. Subjects were asked to record their conversations during a certain period of time, so testifying the actual use of spoken language in accordance with the variation caused by speaker's

parameters. However, in practice, the results are limited to the sole context of chat at home, which is the easiest situation for recording, but provides a reduced variation of speech activities.

It should be clear that providing data through a statistically significant sampling of the population does not imply that all linguistic variations in the corpora are due to the sociolinguistic qualities of the speakers, i.e. age, education, geographical origin, role of the speaker in society (MORENO-FERNÁNDEZ, 2005). For instance, a story told to a child and a row between husband and wife, which are my favorite examples, vary a lot depending on topics, language register, lexical choice and syntactic complexity according to the socio-cultural level of the speakers. However, crucially for pragmatics, the illocutionary quality of the utterances recorded therein can be better foreseen on the basis of context requirements. *Veiled threats*, *protests* and *refusals* will have high probability in a row regardless of the demographic sample. On the other hand, reported speech, narration, explanations have high probability of occurrences in storytelling.

In short, a sociological sampling of the population is valid in so far as it also captures relevant context variations, which is highly predictive of illocutionary variation. Therefore the sampling strategy must capture a huge amount of context variation in order to be a valid source of data for pragmatics. Assuming this conclusion, the comparison with lexical frequency corpus sampling needs can help us understand what the guidelines for setting up a valid corpus for pragmatics study should be. To the ends of lexical frequency the variation in topics and the wording actually used can be derived from a higher or lower probability of occurrence of those arguments in the world. The sampling is adequate when all parameters have probability of occurrence in their relative frequency. This need is much less relevant for speech act types, for which the goal is not to retrieve the relative frequency of a type, but rather, at present, to identify most possible types and their qualities. For this reason the contextual variation testified in the corpus must focus on variation of contexts rather than on their probability of occurrence.

4. Corpus data and the definition of *speech act* types

4.1 The illocutionary values of intonation

The definition of the value of an utterance as a conventional activity performed in the speech flow is strongly dependent on interpretations that may

be highly subjective. It can reach a sufficient degree of inter-rater agreement when the appropriate tagset for a well-defined situation is applied, as is the case with map task (CARLETTA *et al.*, 1997), but remains vague in an unlimited context.

For instance the following dialogic turn has been interpreted within the LABLITA tagset as a sequence of one question, an alternative question; a self-answer; and one act of conclusion (tags in the annotation line %ill).

*SUS: [1] *lei /gliene serve una anch'a lei ? [2] una in più / o no ? [3] no //* [4] *lei ha questa //* [you / (do) you need one also for you ? one more / or not ? no // you have this one //] %ill: [1] question; [2] alternative question; [3] self-answer; [4] conclusion

This annotation has been done mainly on the basis of the interpretation of the value conveyed by the prosody of each utterance and the value of the semantic content in that context. As a matter of fact, in corpus-based research relevant *speech acts* are not identified either on the basis of the occurrence of performative verbs or on the basis of the logic of conversation, as in the Searle/ Grice paradigm (SEARLE, 1975, GRICE, 1975).

Let's concentrate on the fourth utterance that has been tagged as a "conclusion". It must be highlighted that, on the basis of possible performative paraphrases and contextual adequacy, the value *verification* could also have been assigned, or alternatively the simple value *assertion*. This question could be considered underdetermined in principle. We will show in the following that it is not undeterminate, if the differential value of prosody is considered. We will see that the exploitation of prosodic cues is crucial if spoken corpora must contribute to pragmatics.

In ordinary speech, prosody is essential to pragmatic interpretation. In natural speech a language string cannot receive an interpretation at all without prosody, which is the necessary interface between the illocutionary and the locutionary act. This is obvious when speech acts like *assertions, orders* or *questions* are concerned. It is well known that every language has melodic shapes conventionally codified to express sentence modalities, and this is one of the main functions of prosody. For instance, the following Italian phrasing (*gira a destra* [turn right]) can perform in a given context either one *order* or one *assertion* according to its prosodic form.
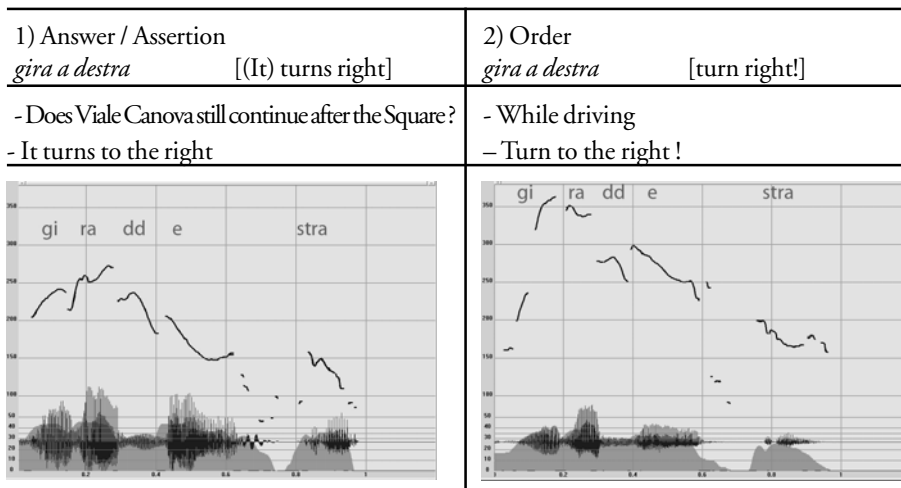
1) Answer / Assertion	2) Order
gira a destra [(It) turns right]	*gira a destra* [turn right!]
- Does Viale Canova still continue after the Square? - It turns to the right	- While driving – Turn to the right !

FIGURE 2 - Answer vs. Order[4]

Although the theoretical framework for the description of prosodic features may vary, it can be verified that the prosodic form of the two acts have differential features. The following graph shows the two previous F0 curves overlapped in transparency.
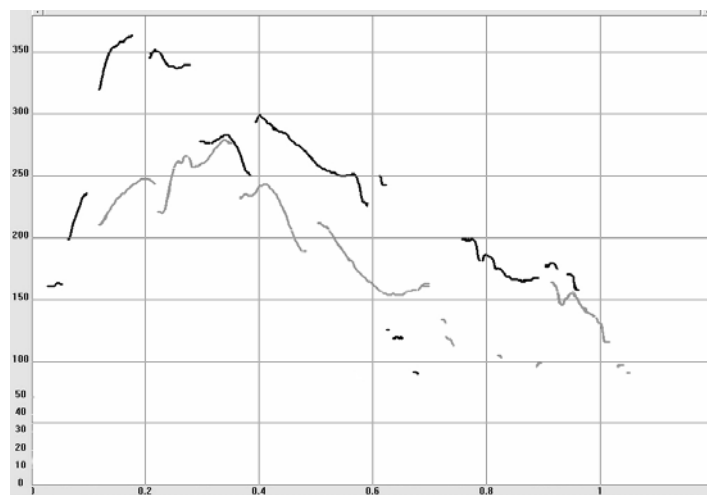

FIGURE 3 - Answer (gray) vs. Order (black) overlapped in transparency

[4] These and the following graphs have been generate by the speech software WINPITCH-PRO and correspond to the same female voice.

Very roughly speaking the prosodic nucleus of an assertive utterance (in gray) is characterized by:

Rising-falling movement. Rising at a mid F0 value followed by a gradual falling
- the post tonic syllable is longer
- mid intensity

The nucleus of an order (in black) is characterized by:

- rising-falling movement. Short optional rising preparation followed by a rapid falling (tail) on the tonic syllable, starting at high F0 values
- the post tonic syllable is short
- high intensity

Most scholars will also agree that the above profiles have a differential pragmatic value; i.e. they are a necessary feature in order for the utterance to be interpreted as an order or an assertion. This can be easily verified in an experimental setting. Given a pragmatic context requiring an order, the replacement of the appropriate order with the same stretch of speech intonated with an assertive profile is meaningless. We have carried this experiment out and the result is impressive (see below).

However, it is much less obvious to what extent the relation between prosody and speech acts characterizes ordinary speech and to what extent the study of prosodic profiles retrieved in spoken corpora can really help to characterize the system governing speech act performance. As we have just observed in the previous example, an *assertive* act can, in principle, be interpreted as a *conclusion or alternatively as a verification* and the potential adequacy in the context of an equivalent performative sentence ("I conclude that ..." "I verify that …") does not really select the actual interpretation. As a matter of fact, we do not have explicit criteria to distinguish an "assertion" from a "conclusion" in the set of utterances which commit the speaker with the truth.

Moreover, sometimes the label derived from the interpretation of corpus data is not a performative verb. For instance, *to instruct* is not a performative verb, but the language activity to give *instruction* has been retrieved in all previous corpus studies as such. What ensures that this activity corresponds to an illocutionary act? Can we set up the conditions determining that an *instruction* is performed rather than just an *order* or a generic directive act? A tagset of 90 labels for speech act types needs very detailed specifications

in order to be applied; otherwise, the definition of types within each illocutionary class remains underdetermined.

4.2 Prosody and "Empirical" Pragmatics

In this section the methodology for studying how prosody contributes to the illocutionary interpretation of spoken utterances will be sketched. To this end the interpretations derived from corpus analysis must be challenged on empirical grounds. The prosodic profile of the utterance to which the tag has been assigned must be repeated with different locutive content by different speakers. The appropriate context eliciting the act can be defined. The following is a summary of the standard empirical methodology for the exploitation of corpus data to the ends of empirical pragmatics:

Corpus driven induction
- collection of speech acts occurrences that have been judged to be of the same illocutionary type during corpus annotation

Positive repetition of the profile in controlled elicitation context
- operative description of the pragmatic characters of the elicitation context
- production and validation of a fictional elicitation context for one comment with the appropriate profile
- repetition and validation by different locutors of the profile as a function of the elicitation context
- adjustment and definition of the pragmatic features of the elicitation context that better allow the production of the profile

Substitution test
- the stretch of speech performed with the appropriate profile is substituted by the stretch of speech with other profiles
- competent speakers evaluate whether or not the profile fits the circumstances

Differential prosodic properties
- repetition of the profile on different accentual structures in the elicitation context
- description of the differential prosodic properties of the profiles
- synthetic modification of necessary features and validation of the range of accepted modification in the eliciting situation (not discussed in this paper).

Below, the overall problem of what determines the interpretation of a speech act in ordinary speech will be grounded, through questioning whether

or not we can find differential prosodic features between the two acts studied, respectively *assertion* vs. *conclusion* and *order* vs. *instruction*.

Figure 4 presents the profiles found in our corpus in utterances respectively tagged as *answer, conclusion, instruction* and *order* that have been repeated by the same female locutor in experimental setting with respect to the same Italian locutive content "Gira a destra" [turn on the right (one)] which has been chosen for its semantic ambiguity . Figure 5 presents the overlapping transparencies of the two pairs of curves:
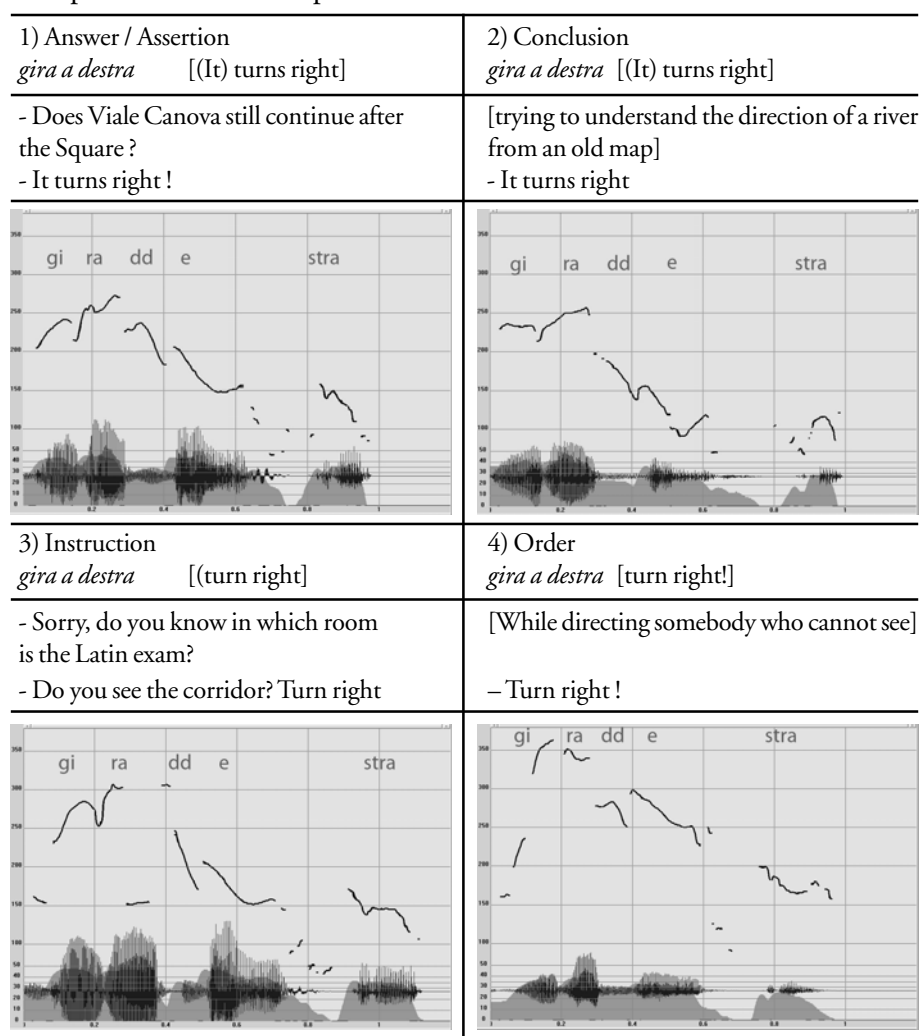
1) Answer / Assertion *gira a destra*　　　[(It) turns right]	2) Conclusion *gira a destra*　[(It) turns right]
- Does Viale Canova still continue after the Square ? - It turns right !	[trying to understand the direction of a river from an old map] - It turns right
3) Instruction *gira a destra*　　　[(turn right]	4) Order *gira a destra*　[turn right!]
- Sorry, do you know in which room is the Latin exam? - Do you see the corridor? Turn right	[While directing somebody who cannot see] – Turn right !

FIGURE 4 - "*gira a destra*" [(0 turns right] with the prosodic profiles of Anwer (1); Conclusion (2); Instruction (3); Order (4)

As can be seen from the overlapping of the two pairs of acts, the F_0 profiles of each speech act type belonging to the same illocutionary class are quite different.

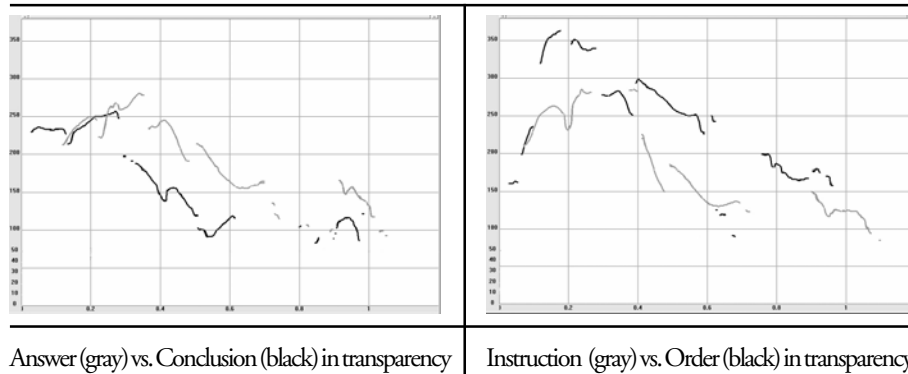| Answer (gray) vs. Conclusion (black) in transparency | Instruction (gray) vs. Order (black) in transparency |

FIGURE 5 - F_0 Illocutionary acts of the same illocutionary class in overlapping transparencies

The table below reports the main prosodic characteristics (with regard to F_0 and Duration properties) that have been highlighted in the study of the profiles in accordance with the IPO system.[5]

---

[5] The movements of the nucleus of the tone unit are described according to the IPO terminology in 't Hart, Collier, & Cohen 1990. The following is the matrix of possible movements.

	/1/	/2/	/3/	/4/	/5/	/A/	/B/	/C/	/D/	/E/
rise	+	+	+	+	+	-	-	-	-	-
early	+	-	-	-	+	-	+	-	-	+
late	-	+	-	+	-	-	-	+	+	-
spread	-	-	-	+	-	-	-	-	+	-
full	+	+	+	+	-	+	+	+	+	-

The system works according to the following parameters: Direction of Movement (rise-fall) / Position of Movement in the syllable (early-late) / Duration on syllables (spread or not) / If the movement cover or not the maximum-minimum excursion (full or not). See Firenzuoli (2003) for a more comprehensive framework.

TABLE 3
F_o Features

Illocution	Onset	Level	Range	Nucleus	Structure	Alignment
Answer	Mid/Low	Mid	Mid 100/200 Hz male 200/300 Hz female	[1A][D]	(Pre-nucleus) (Tail)	Right side of the tone unit [1A] [D] on the focus
Conclusion	Low	Low	100/150 Hz Male 150/250 Hz Female	[D]	(Pre-nucleus)	Right side of the tone unit
Instruction	Mid	Mid/High	Strong 80/250 Hz male 150/300 Hz female	[1A] [D] (Final Plateau)	(Pre-nucleus)	Left side of the tone unit [1A] Right side [D] (Final Plateau)
Order	Mid/High	High	Strong 150/250 Hz male 100/400 Hz female	[1A] sudden	(Pre-nucleus) (Tail)	Left side of the tone unit [1A]

TABLE 4
Duration Features

Illocution	Syllable Length of the nucleus	Speed of the utterance
Answer	Around 200ms.	Mid
Conclusion	Around 350 ms	Mid
Instruction	Around 250ms.	Slow
Order	Around 150ms.	High

The speech acts with the prosodic profiles recorded in the above tables have been challenged. If a specific prosodic profile constitutes a differential feature in order to attribute the requested illocutionary value in the appropriate pragmatic context, then the illocutionary value is conventionally codified within the language system (reglardless of its lexical performative encoding). This is what is required by the system. The following tables summarise the pragmatic features that are needed to characterize the elicitation contexts for *answer / conclusion / order / instruction.*

TABLE 5

Pragmatic features of the elicitation context for Assertion vs. Conclusion

	Answer	**Conclusion**
communication channel	open	open
attention	shared	shared
prossemic between the speakers	direct interaction	no interaction
intentional features of the process	cognitive	cognitive
effect	shared information focus	indirect information focus
modifications in the partner	cognitive	not implied
perceptual characters of the referred objects in the pragmatic/ cognitive context	no restriction	proximal
preparatory condition in the speaker	question by the hearer	problem in the context
preparatory condition in the hearer	expectation	no restriction

TABLE 6

Pragmatic features of the elicitation context for Order vs. Instruction

	Order	**Instruction**
communication channel	open	open
attention	shared	shared
prossemic between the speakers	direct interaction	direct interaction
intentional features of the process	behavioral	cognitive
effects	modification of the world	modification of knowledge and abilities
modifications in the partner	operative	cognitive
perceptual characters of the referred ontological entity in the pragmatic/cognitive context	presence of the referred ontological entity in the context	possibility to explore the context
preparatory condition in the speaker	social role and/or pragmatic skill	knowledge
preparatory condition in the hearer	possibility of intervention in the pragmatic situation	need of know-how

These features are instantiated in scenes performed by actors and represented in Figure 6. Scenes are eliciting context for the appropriate prosodic profile.
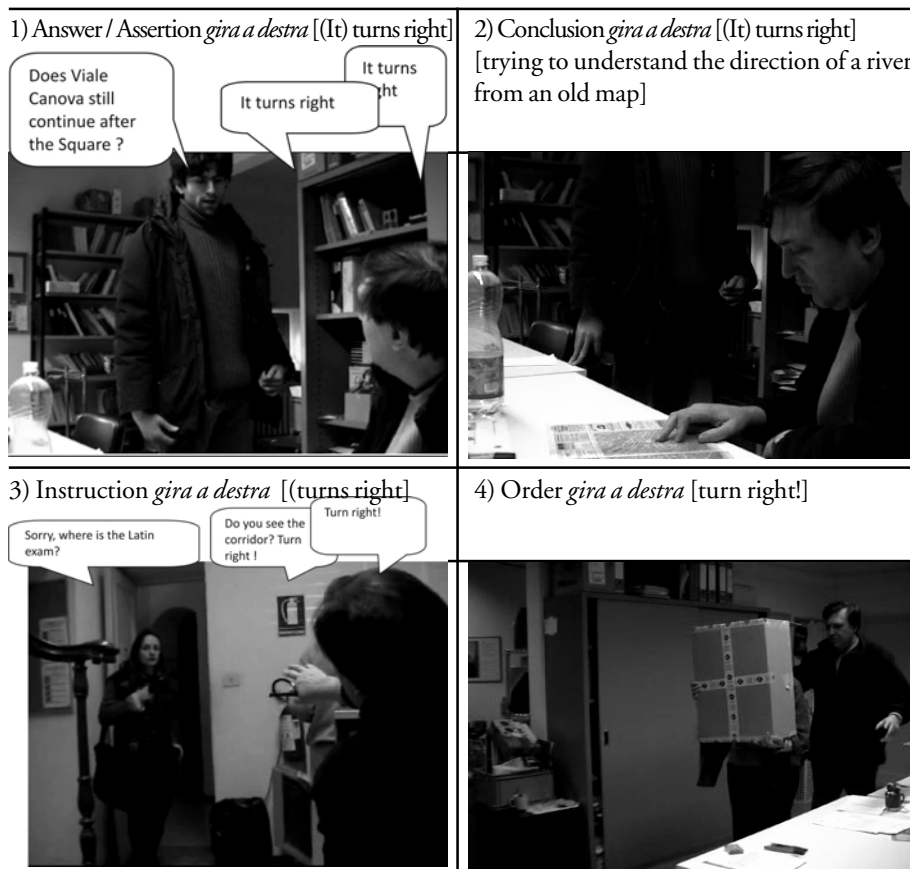
1) Answer / Assertion *gira a destra* [(It) turns right]	2) Conclusion *gira a destra* [(It) turns right] [trying to understand the direction of a river from an old map]

Does Viale Canova still continue after the Square ?

It turns right

It turns right

3) Instruction *gira a destra* [(turns right]	4) Order *gira a destra* [turn right!]

Sorry, where is the Latin exam?

Do you see the corridor? Turn right !

Turn right!

FIGURE 6 - Elicitation contexts for Answer, Conlusion, Instruction, Order

In the elicitation contexts the profiles work fine with the appropriate illocutionary values and are easily replicated by the speakers. Elicitation features are quite compulsory. For instance, in context 6.2, we discovered that as soon as the actor addresses the utterance to the hearer looking at him, despite the intention to replicate the *conclusion profile*, the outcome bears the *answer profile*, while the utterance can be naturally performed with the *conclusion profile* when the speaker does not look at him, but rather concentrates on the object he is evaluating. The feature "no interaction" is therefore a necessary trait for acts of *conclusion*. If the *conclusion profile* is forced in the context eliciting the answer, the speech act is judged as "depressed utterance". If, on the other hand, the *answer profile* is forced in the eliciting context of a conclusion, than the utterance is not judged as a conclusion any more, but rather as a simple assertion.

In the case of Figure 6.4 we discovered that the *order profile* is hard to be replicated in all contexts in which the hearer understands what to do on the basis of his evaluation of the information provided in the order. The *instruction profile* is performed instead of the *order profile*. In context 6.4, this is not the case and the order profile is easily elicited. The differential feature "behavioural" vs. "cognitive" is crucial to foresee if the prosodic profile of "order" or the prosodic profile "instruction" will be performed. When the instruction profile and the order profile are forced in a context that is adequate to the other, the result is totally unacceptable.

We must underscore that features that are responsible for the above systematic prosodic variation bearing illocutionary value; i.e. the underling linguistic form of the speech act, cannot be identified on the basis of any deductive process. In particular, "shared attention without eye contact" for *conclusion* and "Cognitive vs. Operational process in the interlocutor" for *order* and *instruction*, must be considered idiosyncratic constraints (naturalistic) that can display their pragmatic relevance on the basis of an empirical investigation.

The idea that context determines the illocutionary interpretation of utterances, which originated from Austin, does not fit in with the above experimental data. In order to get an utterance appropriate interpretation, a specific prosodic profile is required as a necessary condition. From a logical point of view, we would have expected, for instance, the *instruction*-intonated utterance and the *order* to be both acceptable in the above contexts, since they belong to the *directive* class; but the intonation requirement is compulsory. While the context supports and elicits the prosodic performance of the utterance, it does not determine its value. Therefore we can infer that the distinctions between order vs. instruction and answer vs. conclusion respectively follow from conventional features borne by prosody, and are therefore genuine illocutionary distinctions codified within the language system.

The identification of these speech act types, as many others, is strictly dependent on the availability of large corpus data in which those acts have probability of occurrence. The definition of the pragmatic constraints to the performance of those acts (i.e. the semantic forms underlying them) is not associated to a lexical item but rather to prosodic forms. Therefore, in summary, corpus data enhance pragmatic theory in two main respects, both crucial:

a) The possibility to have a clear picture of the natural speech act types actually performed in ordinary language usage requires corpora covering a huge variety of contexts in which those acts have probability of occurrence and

a long work of annotation and experimental verification; b) the inner semantic form of speech act types can be derived from pragmatic investigations which are grounded on experimental data rather than on sole logical inference.

5. Information patterning and pragmatic functions in the Language into Act Theory

5.1 The *comment* unit

Prosody carries out various functions: a) segmentation of the speech continuum into groups (structural function); b) expression of differential modal acts (statement, order, question etc.) regardless of their segmental content; c) expression of emotions and attitudes (not considered in this paper) (BOLINGER, 1972; 1989; CRYSTAL, 1969; CRYSTAL; QUIRK, 1964; DANEš, 1960; LADD, 1980; ROSSI, 1985; 1999).

The structural function is in principle separated from modal and expressive functions and for what regards corpus data it is linked to the need for annotation of prosodic parsing inside speech transcriptions (BRAZIL, 1995). For instance the annotation of prosodic parsing in the Santa Barbara Corpus of American English (DUBOIS *et al.*, 2000) foresees the marking of both terminal and non terminal breaks. In that coding scheme (CHAFE, 1993) Intonation Units are considered basic organizational units of speech, but according to the overall conception of spoken language (CHAFE, 1987; 1994) they represents *ideas* activated at the consciousness level and bring about the flow of thought rather than speech acts. The previous arguments regarding the importance of terminal breaks for marking speech acts boundaries go in a different direction. In this section we will present other arguments to demonstrates that, besides the segmentation of the speech flow into speech acts, the internal prosodic parsing of the utterance is also relevant to pragmatics if we want to exploit spoken corpora for its ends.

Although the number of utterances without internal prosodic segmentation is consistent in interactive speech, in the majority of cases utterances are not composed of a single word-grouping, but correspond to a complex pattern.[6] From this point of view the utterance has often been

---

[6] In the Italian C-ORAL-ROM subcorpus, the percentage of utterances made up of groupings of more than one word is over 57%, but in the formal domain it is generally much higher (CRESTI; MONEGLIA, 2005).

indicated by a dual functional opposition (WEILL, 1844; MATHESIUS, 1929;) in terms of theme/propos (BALLY, 1950), theme/rheme (Prague functionalism, SORNICOLA; SVOBODA, 1989), topic/comment (HOCKETT, 1958), topic/focus (CHOMSKY, 1971; JACKENDOFF, 1972; LAMBRECHT, 1994), given/new (HALLIDAY, 1976), prefix/noyau (BLANCHE-BENVENISTE, 1987; 2003). More recently a lot of scholars interested in dialogue structure and pragmatics have also focused on other components of the utterance that have a clear pragmatic value, that is, discourse-makers and the functions they carry out to regulate the dialogue (SCHIFFRIN, 1987; BAZZANELLA, 1995; BAZZANELLA *et al.*, 2008).

Although almost all researchers noticed that word grouping is marked by prosody, only few of them have exploited this property to study the pragmatic organization of speech. The research carried out at LABLITA in the frame of the Language Into Act Theory (CRESTI, 1994; 2000; CRESTI; MONEGLIA, 2010) points out that the annotation of prosodic parsing is strictly necessary to specify functional structures of the utterance and specifically those components that have pragmatic values. [7]

In this frame every utterance corresponds to an information pattern which is systematically signalled by an intonation pattern whose units are marked by non-terminal prosodic breaks. The intonation pattern is therefore isomorphic to an information pattern. The most important innovation brought by the Language into act Theory is that information is ruled within actual spoken language use according to pragmatic principles (CRESTI, 1987). The core of the utterance does not correspond to a predication or to a focus, but rather to the expression of the illocutionary value. Crucially the expression of the illocutionary force is up to one and only one word grouping within a prosodic envelope. The information unit so defined (the Comment) accomplishes the illocutionary force and for this reason it is the only unit which is necessary and sufficient to give rise to an utterance.

In other words, in spontaneous interactive speech, if the utterance is simple, i.e. it is made up of one prosodic envelope only, then it does not show an information structure and necessarily bear the prosodic cues for its illocutionary interpretation. If on the contrary the utterance is made up of more then one envelope, than only one of them bears illocutionary cues (the

---

[7] The Informational Patterning Theory has been introduced in Cresti (1987) and Cresti (1994) and developed in many publications after the reference book (CRESTI, 2000). See also the debate on *Macro-syntax* in Scarano (2003).

Comment). For this reason the Comment unit constitutes the core information unit of the utterance, i.e. the utterance cannot be interpreted at all if this unit is erased. On the other hand, in a complex utterance all units other than the Comment can be erased without compromising the possibility of an interpretation to be assigned.

This is a severe constraint regarding the way the illocutionary force of the utterance is expressed in spontaneous speech. Whatever the length of the utterance parsed by prosody might be, there is always one and only one prosodic unit which bears the illocutionary cues allowing its interpretation.

For instance, let us consider the first utterance of the previous dialogic turn which is made up of two prosodic units:

*SUS: [1] *lei /gliene serve una anch'a lei ?COM* [you / (do) you need one also for you ?]

The second unit is the Comment and it is necessary and sufficient for the interpretation. This is not because of its sentential form. For instance, in the following examples, taken from C-ORAL-ROM and C-ORAL-BRASIL, a clause structure appears in the second unit, but this is not the Comment. The first unit can be interpreted in isolation, while the second, although it contains a verb, can be erased without the loss of illocutionary value.

*LIA: Baratti /COM mi pare fosse <stato> // PAR [ifamcv01] [the Baratti Goulf, I think It was that place]
*LUZ: duas vagas /COM eu acho //PAR [bfamdl03] [two positions, I think]

The second unit has a propositional form, but it cannot be interpreted in isolation since it lacks the prosodic information which specifies how to relate it to the world. If the first unit is erased, it is perceived as "suspended". In the following verbless utterance (again, an answer), the illocutionary cues are in the second unit, which, again, cannot be erased.

*LID: il mi' bisnonno /TOP Pietro //COM [ifamdl02] [My grandfather, Peter]
*LAU: departamento /TOP Artes Plásticas //COM [bfamdl03] [Department, Fine Arts]

It may be interesting to note that we cannot provide any distributional evidence for the above assertions without taking into account the prosodic form and the ability by competent speakers to assign or not an interpretation. In other words, the distributional evidence requires both an acoustic source and speaker's judgments.

In light of the above considerations we can underscore the following requirements for pragmatic analysis of spontaneous speech corpora. The internal organization of the utterance into prosodic units is an essential feature of spoken corpora annotation. In order to identify the illocutionary values expressed by prosody, the speech flow must be parsed into terminal and non terminal prosodic units and for each utterance the autonomous unit must be identified by competent speakers. This task highlights the core of the utterance and distinguishes the illocutionary information unit (the Comment) from all other units. The pragmatic definition of the Comment unit within the utterance structure is a crucial finding to bootstrap the illocutionary values conveyed by prosody from corpora. The relevant prosodic cues are foreseen in one unit of the utterance only, whatever the length of the utterance might be, so making this task affordable.

5.2 Speech acts and Dialogue acts

Within the utterance, various types of information units, all optional from an informational point of view, can surround the Comment.[8] They also correspond to prosodic units and are divided into two classes dedicated to different types of functions: a) the textual construction of the utterance (Topic, Appendix, Parenthesis, Locutive Introducer – not considered here); b) its communicative support (Incipit, Phatic, Allocutive, Conative, Connector.[9] These last units are of special interest to pragmatics, and the prosodic annotation of internal prosodic boundaries of the utterance is again essential to their corpus based identification.

In the last twenty years, new data derived from a better consideration of spoken language have made it clear that language expressions proper have not been clearly identified in the grammatical tradition. These expressions, usually referred to with the term "Discourse markers" (SCHIFFRIN, 1987), are dedicated to perform the peculiar pragmatic functions that are required to manage the dialogic interaction. More specifically, they are signals directed to the interlocutor, carrying some specific functions. For instance:

---

[8] The main aspects of the informational patterning of the utterance are described in Cresti (2000).

[9] "Substantial" vs. "Regulatory" Intonation units, in Chafe's terms.

- turn taking
- attention request
- opening of the communication Channel
- phatic function
- keeping the communication channel open
- reception control

Discourse markers are found in all languages and can be roughly identified on the lexical level. For instance, in English expressions like *listen, guys, I mean*; in Spanish *o sea , pues nada*; in French *hein alors donc écoute*; in Italian *senti, guarda, allora, eh*; Brazilian Portuguese *né, cara, oi' ,* may play this role.

Although most of these expressions might be, in many cases, assigned to their traditional Part of Speech, in certain positions of the speech flow, they lose their usual meaning and morphosyntactic value and play a dialogue regulation function instead. Moreover, in conjunction with this peculiar value, these expression are not any more compositional, i.e. they do not contribute to the propositional meaning of the utterance, and can, in principle, be eliminated without effect on the propositional meaning itself (BAZZANELLA, 2006; SCHOURUP, 1999; FRASER, 2006). There is no agreement as to the number of Discourse Markers, their functions, nor criteria for their definition (FISCHER, 2006).

This important chapter of present day's pragmatic theory can strongly benefit from corpus data only if the prosodic and informational properties of discourse markers are taken into account. Indeed, and also in this case most authors have noted the strong correlation with prosody; in particular, discourse markers tend to occur in the dedicated tone unit in which they are isolated.

Indeed, this property fits in with the general framework of the Language into Act Theory. Discourse markers are just information units and, in accordance with a general principle, they show in one-to-one correspondence with prosodic units. Discourse markers fit in with units playing a set of functional roles in the frame of dialogue regulation.

The prosodic character can allow the individuation of discourse markers in speech, i.e. it helps to specify when the above lexical items are compositional elements which play their usual PoS role and conversely when they work as dialogue regulators. For instance in the following two examples the Italian verb *guarda* [look] is used as a discourse marker in a *conative* function (to push the interlocutor to a shared point of view (CRESTI, 2000). It is isolated respectively

in first and final position, it is not compositional and for this reason it is hardly thought of as a verb.

*LIA: guarda / io ti dico così // [look / I will say this way //
 *SRE: ti stavamo aspettando / guarda // [we were waiting for you / look]

The prosodic break is a necessary feature in order to get a Conative dialogue act. It would be unacceptable to group the same stretch of speech within one sole prosodic envelope:

* guarda io ti dico così // [look I will say this way //
* ti stavamo aspettando guarda // [we were waiting for you look]

The opposite occurs when the same expression (*guarda*) is a verb in a functional relation with its propositional object, as in the following example:

*MAX: guarda come tu stavi // [look in what a shape you were]

The expression is not isolated in a distinct prosodic unit and since it is a verb, in no circumstances could it be interpreted as a means to accomplish a dialogue act. The prosodic parsing is the main heuristic to foresee whether the expressions commonly used as discourse markers play a dialogic function or, on the contrary, participate in the construction of the propositional content of the utterance. The frequencies of dialogue-type functions are very high and, more than 50% of utterances performed in spontaneous conversations present these kinds of devices (FROSALI, 2006).

Dialogic units are information units with pragmatic value which on one side must be distinguished from the locutive expressions which contribute to the propositional meaning and on the other from the linguistic units which perform illocutive acts. In other words, from a pragmatic point of view, a clear distinction between speech acts and discourse particles must be made.

The occurrence of a discourse marker in speech may be recorded as a dialogue act in an annotation schema and listed within the series of natural speech acts. This is conceivable since these expressions do not have a propositional value but rather perform an interactive function. The DAMSL coding scheme, for instance, takes this view. However, this is misleading, since these entities are not autonomous from a linguistic point of view, as a speech act should be. In the Language into Act theory these expressions are considered information units within the utterance with a dialogue regulation function, and are clearly distinguished form units which specify their relation to the world (Comment).

For instance Raso & Mello (forthcoming) have distinguished *allocutives* from the *recall illocution* (either proximal or distal). This is an interesting case, since specifically the proper name of the interlocutor can be used in both functions. Indeed grammars consider both under the category of vocative. *Allocutives*, however, are dialogic units of the utterance which play a cohesive and empathic function. They specify to whom the message is directed keeping his attention. *Recall* is a speech act whose function is to get the attention of somebody toward the speaker "opening a close communication channel".

The following examples by the same speaker clearly show their prosodic and functional distinction.

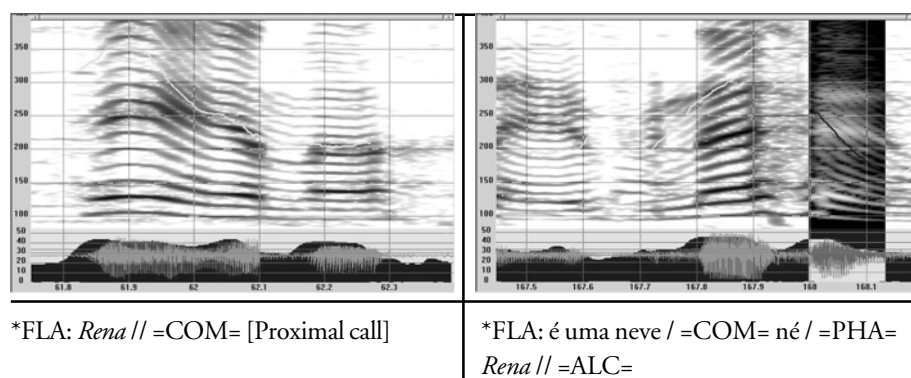| *FLA: *Rena* // =COM= [Proximal call] | *FLA: é uma neve / =COM= né / =PHA= *Rena* // =ALC= |

FIGURE 7 - Proximal call vs. Allocutive Dialogic unit in Brazilian Portuguese

Recall illocutions correspond to a Comment information unit of a one word utterance, showing a higher intensity, a much higher duration and a functional focus that allows for their interpretability in isolation, besides presenting a high F0 variation. On the other hand, Allocutives have a flat or falling profile, without focus, low intensity and duration (roughly 1/5 of the recall) and are one prosodic unit of a complex utterance. The Allocutive cannot be interpreted in isolation. For instance, cutting off the rest of the utterance and keeping "Rena" in the second utterance, the stretch of speech cannot be interpreted as a speech act. On the other hand, if the call is inserted within a stretch of speech both in starting or final position, it gives rise to a terminal break and the stretch of speech is considered a sequence of utterances, i.e. it cannot be integrated within the utterance with another Comment unit.

Conclusions

Our knowledge about the set of possible speech act types is very far from a satisfactory state. The availability of large collections of spoken language corpora is an opportunity for present day pragmatics to bootstrap speech act variation from the actual use of language, so grounding pragmatic knowledge on strong empirical evidence. However corpora compilation and annotation must follows a set of requirements to this end. At the level of compilation, corpora should ensure the maximum context variation to give, in principle, probability of occurrences to all speech act types, whose variation does not depend no either speakers or topics, but rather on what activities are relevant in a given interaction.

The identification of the linguistic counterparts of speech acts in the speech flow (utterances) is the main requirement for what regards corpus annotation. We have shown that the parsing of spontaneous speech into utterances should be a function of prosodic annotation, since prosodic breaks (and terminal breaks in special) are a necessary correlation of the speech act performance. This prosodic parsing has a crucial advantage: it is easily recovered by competent speakers, while both syntactic structure and pragmatic categorization of the speech flow are strongly underdetermined. The access to acoustic information which provides prosodic evidence is, therefore, the basic requirement for whatever exploitation of spoken corpora in the domain of pragmatics.

More generally, empirical pragmatics relies heavily on the study of prosody as far as the exploitation of spontaneous speech data is concerned. The relation prosody/speech acts is crucial for speech act categorization, since in the ordinary use of language, speech acts types are necessarily performed through conventional prosodic forms. More specifically, we have shown that prosody is a differential feature of speech act types, and that those features are strictly required to accomplish the appropriate acts in their elicitation contexts. In short, prosody is the necessary interface between locutive and illocutive acts.

Finally, under the Language into Act theory, we have proposed that the link between prosody and pragmatics also influences the internal information structure of the utterance. Only one prosodic unit within the utterance is devoted to the expression of the illocutionary cues (Comment Unit) and, for this reason, it constitutes the core of the utterance itself. This step allows a clear distinction between the main pragmatic functions performed within the utterance: illocutionary activity performed by the Comment Unit and dialogue regulation activities performed by other units and referred back to the Comment.

References

ABEILLE, A. *Treebanks Building and Using Parsed Corpora*. Dordrecht: Kluwer Academic, 2003.

AMIR, N.; SILBERT-VARODZ, V.; IZRE'EL, S. (2004), Characteristics of intonation unit boundaries in spontaneous spoken Hebrew: Perception and acoustic correlates. *SProSIG*, p. 677-680, 2003.

ANDERSON, A.; BADER, M.; BARD, E.; BOYLE E.; DOHERTY, G.; GARROD, S.; ISARD, S.; KOWTKO, J.; MCALLISTER, J.; MILLER, J.; SOTILLO, C.; THOMPSON, H.; WEINERT, R. The HCRC map task corpus. *Language and Speech,* v. 34, p. 351-366, 1991.

AUSTIN, L. J. *How to Do Things with Words*. Oxford: Oxford University Press, 1962.

BALLY, C. *Linguistique Générale et Linguistique Française*. Berne: Francke Verlag, 1950.

BAZZANELLA, C. I segnali discorsivi. In: RENZI, L.; SALVI, G.; CARDINALETTI, A. (Ed.). *Grande Grammatica di Consultazione*. Bologna: Il Mulino, 1995.

BAZZANELLA, C.; BOSCO, C.; GILI FIVELA, B.; MIECZNIKOWSKI, J.; TINI BRUNOZZI, F. Polifunzionalità dei segnali discorsivi, sviluppo conversazionale e ruolo dei tratti fonetici e fonologici. In: PETTORINO, M.; GIANNINI, A.; VALLONE, M.; SAVY, R. (Ed.). *La comunicazione parlata*. Napoli: Liguori, 2008. v. II.

BERRUTO, G. *Sociolinguistica dell'Italiano Contemporaneo*. Roma: La Nuova Italia Scientifica, 1987.

BIBER, D. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press, 1988.

BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *The Longman Grammar of Spoken and Written English*. London / New York: Longman, 1999.

BLANCHE-BENVENISTE, C. *Approches de la Langue Parlée en Français.* Paris: Ophrys, 1997.

BLANCHE-BENVENISTE, C.; BILGER, M.; ROUGET, Ch.; VAN DEN EYNDE, K.; MERTENS, P. *Le Français Parlé:* Études Grammaticales. Paris: Éditions du C.N.R.S., 1990.

BLANCHE-BENVENISTE, C. Le recouvreman de la syntaxe et de la macro-syntaxe. In: SCARANO, A. (Ed.). *Macro-syntaxe et Pragmatique. L'*analyse Linguistique de l' Oral. Roma: Bulzoni, 2003.

BNC http://www.natcorp.ox.ac.uk/

BRAZIL, D. *A Grammar of Speech.* Oxford: Oxford University Press, 1995.

BOLINGER, D. L. (Ed.). *Intonation*: Selected readings. Harmondsworth: Penguin, 1972.

BUHMANN, J.; CASPERS, J.; VAN HEUVEN, V.; HOEKSTRA, H.; MARTENS, J-P.; SWERTS, M. Annotation of prominent words, prosodic boundaries and segmental lengthening by no-expert transcribers in the spoken Dutch corpus. *In Proceedings of LREC 2002* Paris: ELRA. p 779-785, 2002.

CARLETTA, J.; ISARD, A.; ISARD, S.; KOWTKO, J.; DOHERTY-SNEDDON, G.; ANDERSON, A. HCRC dialogue structure coding manual. HCRC/TR-82. Human Communication Research Centre, University of Edinburgh, 1996.

CARLETTA, J.; ISARD, A.; ISARD, S.; KOWTKO, J.; DOHERTY-SNEDDON, G.; ANDERSON, A. The reliability of a dialogue structure coding scheme. *Computational Linguistics,* v. 23, n. 1, p. 13-31, 1997.

CHAFE, W. Cognitive constraints on information flow. In: TOMLIN, R. (Ed.). *Coherence and grounding in discourse.* Amsterdam: John Benjamins, 1987.

CHAFE, W. (1993). Prosodic and functional units of language. In: EDWARDS, Jane A.; LAMPERT, Martin D. (Ed.). *Talking data:* Transcription and coding methods for language research. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992.

CHAFE, W. *Discourse, consciousness, and time:* The flow and displacement ofconscious experience in speaking and writing. Chicago / London: The University of Chicago Press, 1994.

CHOMSKY, N. Deep Structure, Surface Structure and Semantic Interpretation. STEIMBERG, D.; JACOBOVITS, L. (Ed.). *Semantics*: an Interdisciplinary Reader. Cambridge: Cambridge University Press, 1971.

CRESTI, E. L'articolazione dell'informazione nel parlato. In: AA.VV. *Gli Italiani Parlati:* Sondaggi sopra la Lingua d'oggi. Firenze: Accademia della Crusca, 1987.

CRESTI, E. Information and intonational patterning in Italian. In: FERGUSON, B.; GEZUNDHAJT, H.; MARTIN, P. (Ed.). *Accent, Intonation, et Modéles Phonologiques.* Toronto: Editions Mélodie, 1994.

CRESTI, E. *Corpus di Italiano Parlato*. Firenze: Accademia della Crusca, 2000. v. I-II, CD-ROM.

CRESTI, E.; FIRENZUOLI, V. Illocution and intonational contours in Italian. *Revue Française de Linguistique Appliquée*, v. IV, n. 2, p. 77-98, 2001.

CRESTI, E.; MONEGLIA, M. *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: Benjamins, 2005.

CRESTI, E.; MONEGLIA, M. Informational Patterning Theory and the Corpus based description of Spoken language. The compositiotionality issue in the Topic Comment pattern. In: MONEGLIA, M.; PANUNZI, A. (Ed.). *Bootsrapping Information from Corpora in a Cross Linguistic Perspective*. Firenze: FUP, 2010.

CRESTI, E.; MONEGLIA, M.; TUCCI, I. Annotation de l'entretien avec Anita Musso selon la Théorie de la langue en acte. In: LEFEUVRE, F.; MOLINE, E. (Ed.). *Unités syntaxiques et Unités prosodiques*, *Langue Française*, 2011.

CRYSTAL, D.; QUIRK, R. *Systems of Prosodic and Paralinguistic Features in English*. The Hague: Mouton, 1964.

CRYSTAL, D. *The English Tone of Voice*. London: Edward Arnold, 1975.

DANEŠ, F. Sentence intonation from a functional point of view. *Word*, v. 16, p. 34-55, 1960.

DE MAURO, T.; MANCINI, F.; VEDOVELLI, M.; VOGHERA, M. *Lessico di Frequenza dell'Italiano Parlato*. Milano: ETAS, 1993.

DUBOIS, J. W.; CHAFE, W.; MEYER, C.; THOMPSON, S. A. *Santa Barbara Corpus of Spoken American English – Part 1*. Linguistic Data Consortium, 2000.

FAVA, E. Tipi di atti e tipi di frase. In: RENZI, L.; SALVI, G.; CARDINALETTI, A. (Ed.). *Grande Grammatica Italiana di Consultazione*. Bologna: Il Mulino, 1995.

FIRENZUOLI, V. *Le Forme Intonative di Valore Illocutivo dell'Italiano Parlato:* Analisi Sperimentale di un Corpus di Parlato Spontaneo (LABLITA). 2003. PhD (Thesis) – Università di Firenze, Firenze.

FISCHER, K. (Ed.). *Approaches to discourse particles*. Studies in Pragmatics 1. Bingley, UK: Emerald Group Publishing, 2006.

FRASER, B. Towards a Theory of Discourse Markers. In: FISCHER, K. (Ed.). *Approaches to discourse particles*. Studies in Pragmatics 1. Bingley, UK: Emerald Group Publishing, 2006.

FROSALI, F. Il lessico degli ausili dialogici. In: CRESTI, E. (Ed.). *Prospettive nello studio del lessico italiano* (Atti del IX Congresso SILFI), Firenze: FUP, 2006.

GADET, F. Variabilité, variation, variété. *Journal of French Language Studies*, v. 1, p. 75-98, 1996.

GRICE, H. Logic and Conversation. In: COLE, P.; MORGAN, G. *Speech Acts. Syntax and semantics*. New-York: Academic Press, 1975. v. 3.

HALLIDAY, M.A.K. *System and Function in Language*: Selected Papers. London: Oxford University Press, 1976.

't HART, J.; COLLIER, R.; COHEN, A. *A Perceptual Study on Intonation.* An Experimental Approach to Speech Melody. Cambridge: Cambridge University Press, 1990.

HOCKETT, C. F. *A Course in Modern Linguistics*. New York: The Macmillan Company, 1958.

IMDI http://www.mpi.nl/IMDI/documents/Proposals/IMDI_MetaData_3.0.4.pdf.

IZRE'EL, S. Intonation Units and the Structure of Spontaneous Spoken Language: A View from Hebrew. In: *Proceedings of the IDP05 on Discourse-Prosody Interfaces*, 2005.

IZRE'EL, S.; HARY, B.; RAHAV, G. Designing *CoSIH*: The corpus of spoken Israeli Hebrew. *International Journal of Corpus Linguistics*, v. 6, p. 171-197, 2001.

JACKENDOFF, R. *Semantic Interpretation in Generative Grammar*. Cambridge Mass: MIT Press, 1972.

JURAFSKY, D.; SCHRIBERG, L.; BIASCA, D. Switchboard SWBD-DAMSL Shallow-Discourse-Function-Annotation Coder's Manual, Draft 13. Technical Report TR 97-02. Institute for Cognitive Science, University of Colorado at Boulder, 1997.

KARCEVSKY, S. Sur la phonologie de la phrase. In: *Travaux du Cercle linguistique de Prague* IV, p. 188-228, 1931.

KEMPSON, R. *Semantic Theory*. Cambridge: Cambridge University Press, 1977.

LABOV, W. *The Social Stratification of English in New York City*. Washington D.C.: Center for Applied Linguistics, 1966.

LADD, D. R. *The structure of the Intonational Meaning*. London: Bloomington, 1980.

LAMBRECHT, K. *Information structure and sentence form*. Cambridge: Cambridge University Press, 1994.

LEHISTE, I. The phonetic structure of paragraphs. In: COHEN, A.; NOOTEBOOM, S. (Ed.). *Structure and Process in Speech Perception*. Berlin: Springer-Verlag, 1975.

MATHESIUS, V. La linguistica funzionale. In: SORNICOLA, R.; SVOBODA, A. (Ed.). (1991). *Il campo di tensione. La sintassi della scuola di Praga*. Napoli: Liguori, 1929.

MILLER, J.; WEINERT, R. *Spontaneous Spoken Language*. Oxford: Clarendon Press, 1998.

MONEGLIA, M. The C-ORAL-ROM resource. In: CRESTI, E.; MONEGLIA, M. *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: Benjamins, 2005.

MONEGLIA, M. Units of Analysis of Spontaneous Speech and Speech Variation in a Cross-linguistic Perspective. In: KAWAGUCHI, Y.; ZAIMA, S.; TAKAGAKI, T. (Ed.). S*poken Language Corpus and Linguistics Informatics.* Amsterdam: John Benjamins, 2006.

MONEGLIA, M.; FABBRI, M.; QUAZZA, S.; PANIZZA, A.; DANIELI, M.; GARRIDO, J. M.; SWERTS, M. Evaluation of consensus on the annotation of terminal and non-terminal prosodic breaks in the C-ORAL-ROM corpus. In: E. CRESTI; MONEGLIA, M. (Ed.). *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: John Benjamins, 2005.

MONEGLIA, M.; RASO, T.; MALVESSI-MITTMANN, M.; MELLO, H. Challenging the perceptual relevance of prosodic breaks in multilingual spontaneous speech corpora: C-ORAL-BRASIL / C-ORAL-ROM in Speech Prosody 2010, W1.09, Satellite workshop on Prosodic Prominence: Perceptual, Automatic Identification Chicago. Available at: <http://aune.lpl.univ-aix.fr/~sprosig/sp2010/papers/102010.pdf>.

MORENO FERNANDEZ, F. Corpus of spoken Spanish language – The representativeness Issue. KAWAGUCHI *et al*. (Ed.). *Usage-Based Linguistics Informatics*. Amsterdam: John Benjamins, 2005.

NAKATANI, C.; GROSZ, B.; HIRSCHBERG, J. Discourse structure in spoken language: studies on speech corpora. In: Proc. AAAI-95 Spring Symposium on Empirical Methods in Discourse Interpretation and Generation, 1995.

QUIRK, R.; GREENBAUM, S.; LEECH, G.; SVARTVIK, J. *A Comprehensive Grammar of the English Language*. London / New York: Longman, 1985.

RAPSODIE Project http://rhapsodie.ilpga.fr/wiki/Chaine_de_traitement

RASO, T.; MELLO, H. Allocutives as discourse markers: a comparative corpus-based study for Italian, Spanish, European Portuguese and Brazilian Portuguese. *Proceedings of the 2th International Pragmatics Conference.* Manchester, 3-8 July 2011. Forthcoming.

ROSSI, M. L'intonation et l'organisation de l'énoncé. *Phonetica*, v. 42, p. 135-153, 1985.

ROSSI, M. *L'intonation, le Système du Français*: Description et Modélisation. Paris: Ophrys, 1999.

SCARANO, A. (Ed.). *Macro-syntaxe et Pragmatique.* L'analyse Linguistique de l'Oral. Roma: Bulzoni, 2003.

SCHIFFRIN, D. *Discourse Markers.* Cambridge: Cambridge University Press, 1987.

SCHOURUPS, L. Discourse markers. *Lingua*, v. 107, p. 227-265, 1999.

SEARLE, J. *Speech Acts:* An Essay in the Philosophy of Language. Cambridge: Cambridge University Press, 1969.

SEARLE, J. *Intentionality.* An essay in the Philosophy of the Mind. Cambridge: CUP, 1983.

SEARLE, J. Indirect speech acts. In: COLE, P.; MORGAN, J. L. (Ed.). *Syntax and Semantics, 3:* Speech Acts. New York: Academic Press, 1975.

SHRIBERG, E.; BATES, R.; STOLCKE, A.; TAYLOR, P.; JURAFSKY, D.; RIES, K.; COCCARO, N.; MARTIN, R.; METEER, M.; VAN ESS-DYKEMA, C. Can prosody aid the automatic classification of dialog acts in conversational speech? *Language and Speech*, v. 3-4, p. 443-492, 1998. Special issue on Prosody and Conversation, 41.

SINCLAIR, J. M.; COULTHARD, R. M. *Towards of Analysis of Discourse*: The English Used by Teachers and Pupils. London: Oxford UP, 1975.

SORNICOLA, R.; SVOBODA, A. *Il campo di tensione*. Napoli: Liguori, 1989.

STIRLING, J.; FLETCHER, I.; MUSHIN, R.; WALES, L. Representational issues in annotation: Using the Australian map task corpus to relate prosody and discourse structure. *Speech Communication,* v. 33, p. 113-134, 2001.

SWERTS, M. Prosodic features at discourse boundaries of different strength. *J. Acoust. Soc. Amer.* v. 101, p. 514-521, 1997.

SWERTS, M.; GELUYKENS, R. The prosody of information units in spontaneous monologues. *Phonetica*, v. 50, p. 189-196, 1993.

WEIL, H. 1844. De l'ordre des mots dans les langues anciennes comparées aux langues modernes. In: *The order of words in the ancient languages compared with that of the modern languages translation*, by C.W. SUPER. Amsterdam: Benjamins, 1978.

WINPITCH-PRO http://www.winpitch.com/

YUKI, K.; ABE, K.; LIN, C. Development and assessment of TUFS Dialogue Module-Multilingual and Functional Syllabus. In: KAWAGUCHI *et al. Usage-Based Linguistics Informatics*. Amsterdam: John Benjamins, 2005.