European Science Foundation, Scientific Network “Intersign: Sign Linguistics and data Exchange”
Workschop: “Morphosyntax (text corpora and tagging)”, Certosa di Pontignano, March 12-15 1999

Massimo Moneglia
Linguistic Laboratory of the Department of Italian Studies (LABLITA), University of Florence, Italy
moneglia@unifi.it

A note on spoken language corpora: units of analysis and language sampling strategies

This presentation focuses on two aspects of spoken language corpora which may be relevant in designing sign language corpora and in comparing oral vs. sign languages, namely which are the relevant units of analysis of spoken language and how to design corpora for a correct sampling of spoken language variation. Both questions will be addressed taking as a point of departure the experience of LABLITA in storing and describing child and adult spoken language corpora and the theoretical approach to the study of intonation developed in the laboratory by E. Cresti

1. Utterances vs. clauses

The spoken domain shows different structural differences when compared to that of the written language. Undoubtedly, one of these is the presence of intonation as a necessary component for the realisation of speech. Intonation has a minimum restitution in the conventional writing code, and even if the latter is a derived diamesic system, it is characterised by a textual organisation, sentences based on syntax, which is different to that of speaking. The previous assertion deserves some commentary. In particular the basic linguistic notion useful for the description of spoken language is quite different from that used for written language: the flow of speech is necessarily divided into utterances and the relation of the notion of utterance with syntax and semantics is not obvious. (See Harris 1951; Bar- Hillel 1967; Sornicola 1981; Miller & Weinert, 1998).

In particular, spoken language utterances frequently do not coincide exactly with the structure of a clause, therefore the process of segmentation of speech continuum is frequently arbitrary (Moneglia & Cresti 1997). Example 1 and 2, can make clear the point.[1] In 1) a student address the question whether or not her professor needs a photocopy like other persons in the class. 2) is a dialogue between a two workers which are fixing a car. Both are extremely common empirical situation dealing with spontaneous speech: the flow of speech turns out to be continuous (no pauses) and the process of its partition is radically underdetermined.

The audio files can receive the following rough transcription in the CHAT format. The reader must click on the highlight letters for the multimedia link to the audio files:

Example 1) **Insert link to audio file “spoken1.wav”**
*SUS: lei gliene serve una anche a lei ? una in più o no no lei ha questa //
you do you need one you too ? one more or not no you have this one //
%add: to the professor
%sit: while SUS is going to go for a photocopy

In example 1 a clear question intonation allows to divide the dialogic turn in two parts, however some other segmentation is needed in order to allow the full interpretation of the text, that still

remains obscure. For example in the first segment we do not know the linguistic status of the first pronoun “lei”. Might it be considered a nominal utterance or not ? In the second segment, given that “you have this one” should be a sentence, we do not know the structure of the items “una in più o no no”. In both cases there is no verbs in the surface leading to a possible solution, then the mapping of utterances on clauses structure turns out to be *arbitrary*. Moreover, given that there is no pause in the flow of speech, any definition of utterance like “silent to silent” will lead us to interpret 1) just as one only utterance.

Many people which study spoken language in order to avoid arbitrary decision process, take the view of getting read of such materials. (see. for ex. Biber et alii, 1998) . The words with no clear clause structure should be considered “fragments” like in the square brackets below :

Example 1’)
*SUS: [lei] gliene serve una anche a lei ? [una in più o no no] lei ha questa //
[you] do you need one you too ? [one more or not no] you have this one //

Taking this view more then 40% of spoken language locution turns out to be made up of fragments, that is an extremely relevant quantity of spoken language will be considered out of scientific analysis.

Example two present similar problems. It shows, however, that also the process of separating sentences from the bunch of fragments happen to be arbitrary:

Example 2) **Insert link to the audio file “spoken2.wav”**
*MIC: che macchina l’ è codesta Punto ?
which [kind of] car is that Punto ?

*OPR: Punto milledue mi guardi ?
Punto 1200cc can you check it out for me?

*MIC: i’ ché ti guardo guarda come tu se’ brutto costì sensore temperatura acqua raffreddamento motore
what I schoud check for you look how bad you are there sensor temperature water cooling engine

For example, the first dialogic turn could be parsed in just one utterance with the word “Punto” within the clause. On the contrary it will be parsed in two utterances by considering it as a first clause followed by a nominal utterance. Both are possible clause structures in Italian. The other two dialogic turns present similar problems. In particular the last turns should be made of fragments in almost all its final part, which is long and obscure:

2’)
*MIC: i’ ché ti guardo - guarda come tu se’ brutto [costì sensore temperatura acqua raffreddamento motore]
what I schoud check for you look how bad you are there sensor temperature water cooling engine

From my point of view the basic unit of analysis of spoken language cannot be the concept of clause since is both
too weak and too strong in capturing evidences form spoken language corpora.

2. Intonation and the notion of utterance in spoken language

The problem may be approached on the empirical ground considering that spoken language events have a prosody which envelops each utterance and specifies to our perception the illocutionary force of the utterance itself, namely the communicative value of a single linguistic action (Austin, 1963). From this point of view the textual structure of spoken language turns out to be composed by speech acts whose quality is mainly defined by their intonation. Written language crucially miss both property.

This in not new. Is quite obvious that intonation say something on the action quality of the utterances. At the same time is also obvious that intonation can allow the interpretation even of a single word.

The approach we developed for the analysis of spontaneous speech links the previous general remarks to three arguments regarding intonation and its linguistic role: 1) the idea that utterances are systematically parsed in groups of tone units *(intonation pattern)*, the type of which is discriminated at a perceptive level (’t Hart et alii, 1990); 2) the notion of utterance has prosodic constraints (Crystal, 1975; Halliday, 1976); 3) the functional value of prosody is predictable (Cresti, 1994).

The melodic pattern which scans a possible utterance can be simple, composed of a single tone unit, or complex, in which case it is made up of two or more tone units linked melodically together. To the scanning of an utterance by means of a complex pattern, corresponds the possibility that within it non terminal tone units occur (Pierrehumbert, 1980) for ex the following is a possible patterned declarative sentence in italian:

Example 3) **Insert link to the audio file “spoken 3.wav”**

Carlo / va a Roma // F = Assertion
[Carlo / is going to Rome //]

INSERT here file “IMAGE1” - IN THE TEXT -

On the other hand, an important tradition of intonation studies (Karcevsky, 1931; Crystal, 1975), has always highlighted the fact that there is no such thing as an utterance without a profile of terminal intonation. Considering the question on the theoretical ground the previous classic statement means that we cannot get two utterances in the same prosodic contour. Of course such constraint holds only for the notion of utterance, which *must* have a prosodic counterpart, while does not hold for the notions of clauses and sentences which are independent from prosody.

At the same time it has been noted that, within the possible tone units of an utterance, the tone information which enables one to identify the illocution, or modality, of the utterance lies in a specific tone unit (Martin, 1978). For example in 3) the heat movement in the second unit on the right.

There are interesting consequences from this statement. For ex the following performance of the sentence “Carlo va a Roma” where the two prosodic units have received both a heat movement on the last tonic syllable *must* be considered a sequence of two separate utterances because of prosodic reason. We have a strong perceptual evidence for it that the reader can reply listening to the audio file 4)

Example 4) **Insert link to the audio file spoken4.wav**

Carlo // va a roma //
[Carlo // is going to Rome //]

INSERT HERE file "IMAGE2" – IN THE TEXT-

Given the above, one can deduce that the tone units compounding an utterance must not be considered mere scan units since they have a structural role: tone structure conveys informative values (Bally, 1950; Halliday, 1976). Such value are object of judgements based on the perception of prosodic cues.

By considering such properties, a meaningful generalisation for the analysis of spontaneous speech related to the value of prosody has been obtained. The idea that in each parsed utterance there is always is a particular tone unit having the structural role of expressing the illocutionary value allow the generalisation of this idea in the analysis of spoken texts: it is possible to identify the utterance limit each time the prosody makes it possible to perceive the completion of a speech act (Cresti, 1994; 1996) allowing the pragmatic interpretability of the text.

For example given that we have perceptual evidence that both prosodic units lead to the accomplishment of a speech act in 4), we have also perceptual evidence that only the second unit accomplish a speech act in 3). The first unit cannot be interpreted as an utterance because of its intonation. The reader can easily reply the perceptual result listening in isolation to all segments in the audio files.

It is a general point on spoken language to notice that utterances performed in prosodic patterns have one and only one tone part which contains prosodic cues allowing its pragmatic interpretation also in isolation. We call such units *comment of the utterance*.[2] Such an unit of information is necessary in order to have a speech act. The other possible elements of the prosodic pattern do not have such cues and therefore cannot be interpreted as independent utterance, but necessarily they are part of the linguistic action performed by the comment unit. [3]

In conclusion an *illocutionary criterion*, permits one to perceptively recognise the scan unit which contains the prosodic indices of the linguistic action, and, at the same time it enables one to distinguish such units from the other information units of an utterance, which are optional and can never be pragmatically interpreted. [4] Generally specking we have a method for segmenting the speech continuum in utterances.

The previous relation among elements of a prosodic pattern holds in natural languages like Italian, Franch, or Dutch apart from any consideration regarding syntax. Such relation depends on the action nature of utterances of spoken language and have its surface linguistic signal in prosodic patterning. The following principles summarise the notions useful for a proper identification of the basic units of spoken language:

Interpretation: Illocutionary force specify how to relate a locutive content to the pragmatic context of the utterance

Utterance: is an utterance every meaningful expression (of each type) which can be interpreted in a pragmatic context

Illocutive criterion: prosodic cues specify to perception the accomplishment of a speech act. Therefore speech flow may be segmented into utterances all the times a linguistic action can be judged from prosodic patterning

Comment principle: The illocutionary cues are up to one and only one tone unit in a patterned utterance. According with the illocutive criterion we can discriminate trough perception whether or not a tone unit bears illocutionary information; i.e. whether it can received a pragmatic interpretation or not.

2.2 The basic annotation of prosody in spoken texts

On the basis of the illocutionary criterion the transcription of spoken language must take into account, as a basic level of tagging, both intonation and the segmentation of speech flow into speech acts. The annotation of the utterances and its prosodic parsing is a basic part of the transcription of the spoken language; it only allows one to *see* those sound partitions operated by intonation, which help us to interpret the spoken text. (Moneglia e Cresti, 1997) Such annotation is very simple: for instance simple bar "/" for each not terminal tone unit, double bar "//" to mark the end of an utterance, question mark "?" for end of interrogative utterances, three dots "…" for a suspended utterance. The following signs can be easily implemented in

Terminal tone units // ? ! …

Not terminal tone units /

The previous criteria have been successfully applied to both corpora of adult spontaneous speech and in infant speech (see. Cresti, (in press); Cresti e Moneglia (1993). This has already permitted the signalling of their organisation in *utterances* and that of essential prosodic indices in the transcription of the texts. The following is the transcription with prosodic punctuation and segmentation of utterances in 1 and 2, in accordance with the illocutionary criterion.

Example 1'')
*SUS: lei / gliene serve una anche a lei ? una in più / o no ? no // lei ha questa //
you / do you need one you too ? one more / or not ? no // you have this one //

Example 2'')
*MIC: che macchina l' è / codesta / Punto ?
which [kind of] car is / that one / Punto ?

*OPR: Punto / milledue // mi guardi ?
Punto / 1200cc // can you check it out for me?

*MIC: i' ché ti guardo // guarda come tu se' brutto / costì // sensore temperatura acqua // raffreddamento motore //
what I should check for you // look how bad you are / there // sensor temperature water // cooling engine //

The reader can verify the perceptual relevance of previous segmentation listening in isolation to each prosodic group in the multimedia files. Accordingly to the *illocutionary criterion* and to the *comment principle* only one segment of every patterned utterance will be interpretable as a speech act, while all simple utterances can receive a possible interpretation in isolation. [5]

In conclusion we briefly showed that the partition in utterances of spoken texts relies on a proper analysis of prosodic cues with their functional values rather to the arbitrary process of assigning a syntactic structure to the spoken text. A part from the fact that the text become suddenly interpretable, we just proposed a criterion for segmenting spoken language which have strong consequences on the evaluation of spoken texts, for both qualitative and quantitative analysis. For example there are no fragment in the previous text, that is properly patterned by intonation and divided into utterances allowing full interpretation for both complex and simple prosodic structures. The idea of a huge proportion of fragments in spoken language is the consequence of assigning a clause structure without considering the action value of utterances in spoken language.

The choice of a basic level intonation tagging might appear to be a useless waste compared to different transcription systems of intonation and especially compared to ToBi, which has been frequently proposed as a standard (See Gibbon et alii 1997). Some clarifications seem to be in order. The annotation of prosodic parsing it is not a *transcription* of the intonation as for example ToBi, or MARSEC, in the sense that it does not load external signs, such as numbers or indices or letters onto the text so as to reproduce in a ciphered manner the characteristics of the intonation profiles. Annotation of prosodic parsing is the definition of the prosodic units having functional value not an evaluation of prosody according to o pre-theoretical phonological typology (see. Pierrumbert 1980; Ladd 1996). The demarcation of the utterances and their prosodic parsing according to the illocutionary criterion, which are annotated in the transcription, will constitute the functional correlate of any lather analysis of prosodic cues.

2.3. The correlation *prosody / utterance* is independently motivate

It may be interesting in the context of this conference to underline that the structural link between intonation and utterance as an action notion is strongly based in the process of first language acquisition and therefore, from a theoretical point of view, the previous theory is independently motivated at the language learning level (see Moneglia 1994).

We tried to show in many works that the basic milestone of combinatorial language is the capacity of the child to integrate in a single language action more then one locutive expression and that prosodic cues mark such process in child language(See Moneglia & Cresti 1993 and references in the LABLITA WEB site). In particular in the early transition from one word period to complex utterance each word of child's dialogic turn allow its interpretation as a single speech act. Such interpretability is a function of prosody that signal the action nature of each single word. For example the following dialogic turn is for sure a sequence of distinct acts because of the prosody of each word. Following the *illocutionary criterion* each word can be interpreted in isolation as a separate utterance because of its intonation. The reader can again verify listening in isolation to each segment in the multimedia file

Example 5) **INSERT link to the audio file "spoken5.wav"**

GIU (1;.9.23) lilla // lilla // totta //
clock // clock // broken //
%sit: playng with the clock

Despite the fact that possible cognitive relations and even a predicative structure could be assigned to the sequence of words "clock –broken", they cannot be one utterance as far as the two words perform more then one illocutionary acts. The phase of *illocutionary independent words* in language acquisition appears as an universal character of language learning.

The child in order to rich the stage of combinatorial language must learn some prosodic tool in order to put together two locutive elements in the same act. The child GIU demonstrate such ability in its longitudinal protocols in about one month. The following are two example of the two basic strategy a child can follow in this task[6]: two words in the same prosodic envelope (*linear strategy*) and two words in a prosodic pattern (*informational patterning strategy*):

Example 6) **INSERT link to the audio file "spoken6.wav"**
GIU (2;0.20) quetto / chiae //
this / key //

Example 7) **INSERT link to the audio file "spoken7.wav"**
GIU(2;0.20) mimma // etti mimma //

child // this child

Again the reader can easily verify our interpretation listening in isolation to each word of the multimedia files. The constraint to give an interpretation to each element as a single act has disappeared, while on the contrary following the *comment principle* one and only one of the two tone elements can receive an interpretation in isolation.[7]

Summarising, in the framework we just designed, intonation is not the prosodic interpretation of the syntactic structure of an utterance. By means of intonation language events receive an action value and spoken language basic entities, namely *utterances*, have such an "action nature" expressed by intonation. The detection of prosodic cues, which are object of perception, allows the definition of explicit criteria for the analysis of oral texts. It would be nice comparing sign language, which is an action-language, with verbal language, to offer a proper comparison with spontaneous speech, which we tried to show, is an action language too.

3– Corpus design

3.1. In connection with the study of the specific structure of spoken language it is also a relevant question to identify criteria for setting corpora which can be a good basis for a comparison from the point of view of corpus design. In facts spoken language shows peculiar linguistic properties with respect to its variation at many levels: structure of dialogic situations; sociolinguistic features, goals and regulations of the linguistic process. According to such variation many linguistic properties for spoken language such as MLU, omission and variation of morphemes, word frequency, may vary a lot from text to text [8]. Therefore all such values which are necessary in order to compare spoken language code are a function of the pre-theoretical selections of contexts in corpus design. What texts should better testify spoken language essential properties and which range of variation those properties may have ?

I would like to and up this brief talk presenting the structure of the LABLITA corpus and some consequence that the variation of text may have for the induction from corpora of general character of spontaneous speech

LABLITA, since the beginning of the 70es, has collected a series of corpora of spontaneous spoken language in order to create some databases to study and identify linguistic properties of spontaneous spoken language and in particular its intonation. One of the main issue was to distinguish informal everyday spontaneous speech, like in the above examples, from formal speech. LABLITA corpora are four, transcribed and electronically recorded in CHAT format (Mac Whinney, 1994):[9]

1) An open corpus of spontaneous adult spoken language. 120 texts of spontaneous spoken language of variable length (from two hours to 5 minutes) for about 62h and 583,000 words.

2) Longitudinal corpora of Italian acquisition (575,000 words for 84 hours)

3) Corpus of the cinematographic language transcriptions of 12 significant films in the history of Italian cinema (1948-1994) for 220,000 words.

4) Samples of media language (radio and TV) for 92,000 words

In the LABLITA corpus variation is the result of parameters that are considered significant in a large literature (See. Bilger, 1997; Labov, 1966; Biber,1994; 1988; Berruto, 1987; Gadet; 1996). The first two parameters constitute the structural variation of the corpus: (1) dialogical structure (monologues, dialogues, conversations) (2) social domain of use (family; private life, public life, media productions). The other parameters, namely speaker's age (3), education (4) and job (5) vary in a transversal way in relation to structural parameters showing the diaphasical and diastratical variations of language uses.

The following matrix shows on the vertical axes the variation of the dialogical structure, on the horizontal axes the variation of the social domain of use. A more accurate classification of every horizontal field distinguishes if the texts, all spontaneous, are at some level regulated or not**.**

	FAMILY		PRIVATE		PUBLIC	
	Free	Regulate	Free	Regulate	Free	Regulate
Dialogue Conversation Monologue	20%		35%	15%	10%	20%

Click here for the list of texts of the LABLITA corpus **Insert link to file "corpus"**

3.2. Measures of informal vs. formal spoken language

The LABLITA corpus assure a data base for the distinction between formal and informal speech trough scanning the various domains of spoken language use. For example texts which are "public" and "regulate" frequently will be more formal of texts which are family conversations. At the same time, as the percentage clearly shows, the corpus cover a huge proportion of spontaneous/informal speech with respect to formal speech. In so doing the LABLITA corpus try to highlight the central area of its use, assuring a proper base for induction. Such a choice is essential to rich a better understanding of the specific properties of spoken language and to identify its quantitative limits; that is also important considering that in general we think to language as an "ideal written language", potentially with no limits.

In particular from a quantitative point of view the difference between the basic properties of formal and informal speech is not well known. The following is an example of a formal spoken text (taken during a seminar on child language at the University of Florence). Its transcription was obtained again applying the illocutionary criterion. The difference with 1) and 2) is impressive, but what is about exactly ?

Example 8) **Insert here link to the audio file "spoken8.wav"**
*CEC: brevemente / quello che era stato detto / è che c' era stato / un incremento elevatissimo di protoforme / e

[briefly / what we told / is that it had occurred / a huge increase of proto-morphemes / and

infatti lo si vede dal grafico / a venti mesi / c' è proprio un picco / e / s' era parlato / appunto / sulla /

actually we see it from the graphic / at twenty months / there is right a rise / and / we talked / in facts / about some /

contemporaneità / anche / dell' esplosione lessicale / ma su questo / i dati devono essere riguardati / ancora devo fare ...

contemporaneity / also / with lexical spurt / but about this fact / data must be revised / I still must do …]

Such distintion can be quite easily demonstrated once we get reliable measures of spoken texts not only at the word level but also at the level of its structure. From this point of view the approach we have briefly presented leads to immediate results. The demarcation of the

utterances and their prosodic patterning can receive an automatic verification giving rise to a series of *basic speech measurements*. Measurements are based on concepts of dialogic turn, informative vs. non-informative words, tone unit, utterances. The following is a simplified version which take into account four parameters.

- MLT average length of speaking turn (utterances per turn)
- average length of utterance (MLU- words per utterance and MTU tone units per utterance)
- average length of tone unit (MLTU - words per tone units).

On the basis of the previous indices we get a quantitative description of the basic structure of the single texts. By comparing these numerical data it is possible to define a first classification of spontaneous texts where the distinction between formal and informal speech may be better appreciate. We has verified (Tizzanini 1999) the descriptive capacity of the measurement on a sub-corpus of six texts, four of which (a family conversation, a country wake, a conversation between work colleagues and a conversation between university students) brought together in Group A, are marked by a strong degree of spontaneity. The other two texts (a university lecture and a radio interview) are formal texts (Group B).The following table summarises some of the results obtained:

	MLT (utter./turn)	MLU (words / utter.)	MTU (tone unit/utter.)	MLTU (words/tone unit)
Group A	2.24	6.03	1.98	2.84
Group B	4.04	11.53	4.22	2.58

The quantitative difference between the two groups is immediately obvious. The most spontaneous texts (A) are generally characterised by quick exchange turns (sign / signal of a very tight / restricted dialogue exchange) and short utterances (sign of a simple informative structure). On the contrary, the Group B texts present longer exchange turns and more complex utterances both at the word level (MLU) and at prosodic level (see Cresti & Scarano 1998) (these factors are highlighted by the values which are almost twice those of the Group A texts).[10] So example 7 is an instance of this tendency: just one utterance, compound of 53 words dividend in 17 tone groups.

I would like to and the talk stressing that such an easy result strictly depends on the choice of the basic unit of analysis of spoken language we adopted : the action notion of utterance. No such result can be obtained describing the structure of the previous spoken texts trough the concepts of sentence or clause: 7) is made up of six/seven sentences and their length is irrelevant as a measure of spoken text.

REFERENCES

Austin, L. J. (1962) *How to do things with words*, Oxford University Press, Oxford.

Bacelar Nascimento M.F. (1998) *Quelques resources linguistiques de base pour l'étude des variétés du portugais contemporain*, Revue Française de Linguistique Appliquée, Vol. III – 1.

Bally C. (1950), *Linguistique générale et linguistique française*, Francke Verlag, Berne.

Berrendonner A. et alii (1983), *Principes de grammaire polyectale*, Presses Universitaires, Lyon.

Berruto G. (1987), *Sociolinguistica dell'italiano contemporaneo*, NIS, Firenze

Bettencourt Gonçalves, J. (to be published) *Português Falado: variedades geográficas e sociais* em Actas do Colóquio Internacional "A investigação do Português em África, Ásia, América e Europa", Instituto Ibero-Americano, Berlim.
Biber D. (1988),*Variation across speech and writing*, Cambridge University Press, Cambridge.
Biber D. (1994), *Rapresentativeness in corpus design*, in Linguistica computazionale IX-X.
Biber D., et alii (eds) (1998), *Corpus linguistics: investigating language structure and use*, Cambridge CUP.
Biber D., et alii (eds) (1999), *The Longman grammar of spoken and written english*, London, Longman.
Bilger M. (1997), *Corpus de portugais et d'espagnol*, in Revue de l'Association Français de linguistique appliquée, 2.
Blanche- Benveniste C. (ed) (to be published), Corpus du Français parlé. Echantillonages. Champion, Paris
Blanche-Benveniste C. (1997), *Approches de la langue parlée Français*, Ophrys, Paris.
Blanche-Benveniste C. Bilger M., Rouget c. Einde van den K. (1990) *Le Francais parlé*, Editions du CNRS, Paris.
Bortolini U. & Pizzuto E. (eds.) (1997), *Il progetto Childes Italia*, Del cerro, Pisa.
Cresti E. (1994), Information and intonational patterning in Italian, in B. Ferguson, H. Gezundhejt, Ph. Martin (eds), *Accent, intonation, et modéles phonologiques*, Editions Mélodie, Toronto pp. 99-140
Cresti E. (1996), Speech acts units and informational units, in E. Fava (ed) *Speech acts and linguistic research*, Nemo Padova, pp. 89-107.
Cresti E. (1998), *Critère illocutoire et articulation informative*, in M. Bilger (ed), Acts du colloque Questions de méthode dans la linguistique sur corpus, Perpignan, Champion, Paris.
Cresti E. . (to be published), Corpus di italiano parlato. Campioni, Accademia della Crusca, Firenze
Cresti E. & Moneglia M. (eds)(1993), *Ricerche sull'acquisizione dell'italiano*, Bulzoni, Roma
Cresti E. Scarano A (1998) Sur la notion de Le parlé spontané, in in M. Bilger (ed), Acts du colloque Questions de méthode dans la linguistique sur corpus, Perpignan, Champion, Paris
Crystal D. (1975), *The English tone of voice*, Edward Arnold, London.
Crystal D. (1985), Prosodic Development, in Fletcher P. e Garman M., *Language Acquisition*, Second edition, Cambridge, Cambridge University Press.
De Mauro T. Mancini F. Vedovelli M. Voghera M. (1993) *Lessico di frequenza dell'italiano parlato*, Etass Milano.
Dictionnaire de Linguistique (1973), Larousse, Paris.
Fasold, R., (1990) *The sociolinguistics of Language*, Oxford.
Firenzuoli V. (1999), *Correlazioni e criteri di variaizone tra articolazione dell'informazione e morfosintassi in un corpus di italiano parlato*, submitted to Romanische Forschungen.
G. Berruto (1987), *Sociolinguistica dell'italiano contemporaneo*, NIS, Firenze
Gadet F. (1996) *Variabilité, variation, varieté: le Français d'Europe*, in French Language Studies, 6.
Gadet F. (1989), *Le Français ordinaire*, Armand Colin, Paris.
Gibbon, More, Winski (eds.) (1997) *The handbook of Standards and Resources for Spoken language Systems*, Mouton de Gruyter, Berlin. Halliday M.A.K. (1976) Sistem and function in Language, (Kress. G. Ed) Oxford
Harris, Z (1951) Methods in structural linguistics, Chicago
Hokett,Ch (1963) the problem of universals in Language in Greeberg (ed) Universals in Language, Cambridge Mass 1-29

Hymes D. (1962), *The Ethnography of Speaking*, in T. Gladwin, W. C. Sturtevant (a cura di), *Anthropology and human behaviour*, Anthropological Society of Washington, Washington D.C.
Labov W., (1966) *The social stratification of English in New York City*, Washington D.C., 1966.
Ladd R.D. (1996), *Intonational Phonology*, Cambridge University Press, Cambridge.

MacWhinney B. (1994) *The CHILDES project: tools for analyzing talk*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
Martin Ph. (1978) Questions de phonosyntaxe et de phonosémantique en Française, in *Linguisticae Investigationes*, 2, pp 93-126.
Miller, J e Weinert R., (1998) *Spontaneous spoken Language: Syntax and Discourse*, Claredon Press, Oxford.
.
Moneglia M. (1994b), Intonation, the property of distantiation and the mastering of complex utterance in Italian, *Quaderni del Dipartimento di linguistica*, Università di Firenze, V.
Moneglia M.(1994) The ontogenetic foundation of informational patterning. In B. G.Ferguson, H. Gezundhajt, P. Martin (eds.), *Stress, intonation and phonological models* pp.155-186.
Moneglia M., Cresti E. (1993) "Formazione dell'atto linguistico complesso e intonazione: l'ontogenesi delle relazioni informative in italiano", Cresti E., Moneglia M. (eds.), *Ricerche sull'acquisizione dell'italiano*, Roma, Bulzoni.
Moneglia M. Cresti E. (1997), Intonazione e criteri di trascrizione del parlarto, in U. Bortolini E. Pizzuto (eds), Il progetto CHILDES Italia, Del Cerro Pisa, pp. 57-90
Nespor M. Vogel I. (1986), *Prosodic Phonology*, Foris, Dordrecth.
Pierrehumbert J.(1980) *The Phonology and Phonetics of English intonation*, Cambridge, MIT Press.
Sornicola R. (1981), *Sul Parlato*, Il Mulino Bologna.
't Hart J., Collier R., Cohen A. (1990), *A perceptual study of intonation*, CUP, Cambridge.
Tamburini G. (1994), *Tone units and local F_0 phenomena*, in B. Fergusson, H. Gezundhajt, P. Martin (a cura di), *Stress, intonation and phonological models*, Editions Melodie, Toronto.
Tizzanini, G.(1999) Articolazione dell'informazione in corpus di italiano parlato. Unpublished master thesis LABLITA University of Florence.
Voghera M. (1992), *Sintassi e intonazione dell'italiano parlato*, Il Mulino, Bologna

---

[1] Examples are taken from two texts of the LABLITA corpus of adult spoken Italian see below for details
[2] The term introduced by Hockett, receives here a prosodic definition (See Hockett, 1963)
[3] We are not going to discuss here the role of all units of information that are not a comment unit of information . The informational relation *topic comment* (Hockett, 1963) is also conveyed by intonation. See for this notion the Theory of informational patterning (Cresti 1996; Tamburini 1994).
[4] The results obtained on the basis of the application of the illocutionary criterion are crucially confirmed in the macro-syntactic theory (Blanche-Benveniste 1990; 1998; Berendonner, 1983), for which in spoken language the syntactic "noyau" coincides with the tone unit having illocutionary value.
[5] The first "*lei*" in 1) and the word "punto" in the second dialogic turn of 2) should be considered a *topic* unit of information in the Theory of informational patterning and, by necessity, cannot receive a pragmatic interpretation as a speech act.
[6] We verified those strategies in a large series of longitudinal studies. See child language corpora of LABLITA listed in http://lablita.dit.unifi.it
[7] Notice that the comment principle apply also to linear utterance in early language acquisition, while it is cannot be the case in adult language. That is the main transitional character in the acquisition of prosody (See Moneglia, 1994)

[8] The problem of designing a reference corpus for spoken language is discussed with respect to present corpora of spoken Italian, in particular the LIP corpus, and the LABLITA corpus (see De Mauro et al 1993; Cresti (in press).

[9] These corpora probably constitute, altogether, the biggest database presently available on spoken Italian. For all the corpora the audio is available on cassettes DAT, audiocassettes, magnetic band. The sound source of spontaneous adult spoken Italian corpus is stored in CD in wav format. The acoustic quality is the one which is possible in environmental recordings which are often concealed. Corpora are accessible for scientific purpose on explicit agreement conditions. (see Moneglia, 1998 and the web site http://lablita.dit.unifi.it for a detailed description)

[10] It is worth noting that the average length of the tone unit remains more or less constant in all the texts. Therefore, it is easy to deduce that the tone unit length is due, not so much to informative needs as, to natural execution needs.