Martina Urbani[1]

INSTRUMENTS AND METHODS FOR THE ANALYSIS OF PROSODY

1. Introduction

One of the greatest advantages of using computer-assisted pronunciation and intonation tutors is that the computer serves both as a medium of instruction and as a tool of research (Chun 1998: 66). After recording and digitizing speech, it is fundamental to correctly interpret the waveforms because a good analysis produces a valuable feedback. According to Anderson-Hsieh (1994: 6),

> it has been found that suprasegmentals can be most effectively taught through the use of equipment which extracts pitch and intensity from the speech signal and presents the information on a video screen in real time, providing instantaneous visual feedback on stress, rhythm, and intonation.

As far as language learning is concerned,

> the major benefit of electronic visual feedback for teaching suprasegmentals is that it provides the students with an accurate visual representation of supersegmentals in real time paired with the normal auditory feedback that occurs during speech (Anderson-Hsieh 1992: 61).

As a first step, the visual feedback deals with the immediate visualization of the waveforms which permit us to estimate how high or low, smooth or interrupted they appear. The relative shortness or length of certain waveforms has to do with the duration of syllables. As a second step, it is possible to accurately analyze the data collected in order to examine and compare the annotations of pitches, frequency, duration, and stress.

2. Praat: a speech analysis software

Out of the great number of speech analysis software available, Praat is by far one of the most used systems 'for doing phonetics by computer' (Boersma and Weenink 2007). However, there is 'a wide variety of signal analysis software with features for quick and accurate extraction of frequency, pitch contours, intensity levels, as well as on-screen display of speech sound waves and spectrograms' (Busà 2008: 170). Among the software available are: Visi-Pitch, WinCECIL© (*Computerized Extraction of Components of Intonation in Lan-*

[1] Martina Urbani is a Ph.D. student in English Studies and Applied Linguistics at the University of Padova, Italy. Her current research interests are: intonation, prosody, second language learning, pragmatics and intercultural communicative competence. Università degli Studi di Padova, Italy; email: martina.urbani@studenti.unipd.it.

Linguistics Applied
vol. 4 (2011), pp. 104-115
Martina Urbani, *Instruments and methods for the analysis of prosody*

guage), Sygnalize©, Speech Analyzer©, MSL (Micro Speech Lab), and Emu.[2] Many speech analysis programs 'have become increasingly more accessible in the form of sound and speech digitizers, pitch trackers to produce displays of intonation curves and computer-assisted language learning (CALL) software, including pronunciation tutors, with audio and graphic components' (Chun 2002: 120).

Prosodic analysis is enhanced by the use of Praat. This computer program was developed by Paul Boersma and David Weenink of the Institute of Phonetic Science, University of Amsterdam. Praat analyzes periodical and non-periodical sounds by providing detailed information about the frequency (highness of pitches), duration (tempo) and intensity (loudness) of sounds (Halliday and Greaves 2008: 18). Such information is illustrated in Praat by sound waves which represent duration (in ms) on the horizontal X axis and frequency (in Hz) on the vertical Y axis. The sound waves are also portrayed as a less or more fragmented curve representing the pitch line. Praat accepts a wide variety of sound formats and it has a clear and articulate interface. This speech analysis program is so sophisticated that 'the Praat segmentation, labeling and in particular the signal manipulation and re-synthesis facilities are unquestionably superior to any corresponding functionality that currently exists' (Harrington et al. 2003: 357). Despite its 'unquestionable superiority', this program is not as complex as it could be. The Praat interface is simple and intuitive and it can be easily exploited not only by scholars and experts but also by fledgling students, who download Praat for free and use it to produce displays of intonation curves and to exploit audio and graphic components. Given the extrinsic worth of this software, it is fundamental to learn how to use it properly.

2.1. Recording systems and methods

Since pitch-tracking systems are continuously upgraded and updated, to make a comparison between particular software is not an easy task: while comparing the different systems for making recording the author may already be out of date. However, Ladefoged (2003: 17) listed four properties that make a program a good recording system: (1) a good frequency response; (2) a good signal/noise ratio; (3) reliability and user-friendliness; and (4) the possibility to use and modify the recordings for a long time. These qualities make a computer program reliable and efficient.

Ladefoged (2003: 85-86) made a detailed comparison of the differences between Praat and Macquirer (another pitch-tracking program) and he pointed out that Praat is more accurate. However, he asserted that

> 'the moral of this comparison between two systems of pitch analysis is not that one is better than the other. Both of them [i.e. Praat and Macquirer] provide generally adequate representations of the pitch, and both of them make mistakes. The point to be emphasized is that making a pitch analysis requires careful adjustments of the options' (Ladefoged, 2003: 86).

For this reason, the main point of a correct analysis is to adjust accurately the setting in order to get the best possible representation of the curve.

[2] These software systems are available online and some of them have the advantage of being freely downloadable. It is possible to find Praat at http://www.fon.hum.uva.nl/praat and Speech Analyzer© at http://www.sil.org/computing/catalog/speechanlyzer.html. CSRE (Canadian Speech Research Environment) is available at http://www.icis.on.ca/homepages/avaaz. WinPitch, developed by Philip Martin, is available at http://www.winpitch.com.

First of all, it is necessary to adjust the sound wave settings according to the specific needs and features of the analysis and then one can proceed with the recording process. After recording a sound, one can visualize a sound wave and interrogate it in order to obtain useful information about the properties of the sound displayed on the screen. When the utterance is digitized and pitch-tracked, it is possible to analyze the sound wave of that utterance, and, if necessary, to compare it with other sound waves that other speakers may have produced. This has been used also in second language teaching to provide learners with feedback through the comparison of their productions with NS productions.

To begin with the recording process, one has to predetermine the settings of the recording. The sound can be recorded by two channels: the mono or the stereo. The sampling frequency can be adjusted between 8000 Hz and 192000 Hz. The recording interface is simple and functional, as you can see in the figure (1).

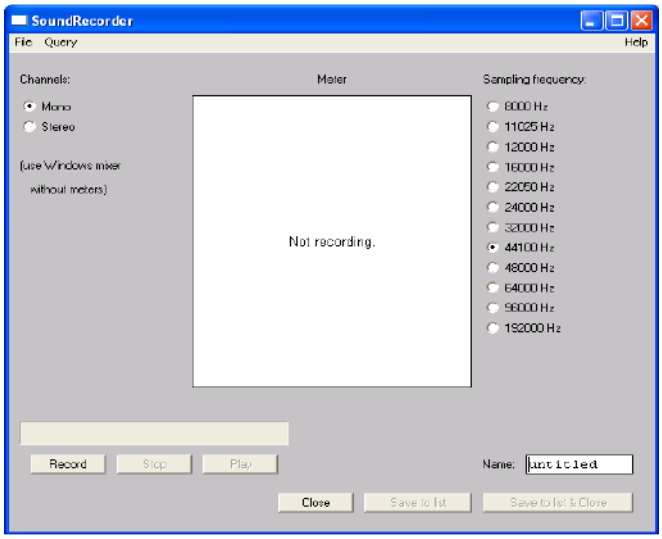


Figure 1. *Sound Recorder Interface.*

After recording an utterance and save it in the saving list, the utterance is processed and visualized in a spectrogram. This spectrogram can be of two types: narrow-band or wide-band. The fist one is very accurate in the representation of time dimension, while the second type is more precise at representing the frequency dimension (Ladefoged 2006: 202).

2.2. Pitch analysis and variations

Just as the spectrograms are visualized according to parameters of view range (Hz), window length (s) and dynamic range (dB), pitch settings influence the display of pitch variations. Whenever a sound is recorded, it assumes different values according to the pitch settings. In the pitch setting window, it is possible to adjust the pitch range, the pitch unit and the drawing method. As you can see in the figure below, pitches are measured in Hz with an automatic drawing method. As for the pitch range, it depends on the quality of the sounds that are going to be recorded and on the kind of visualization that can be more suitable for the analysis.

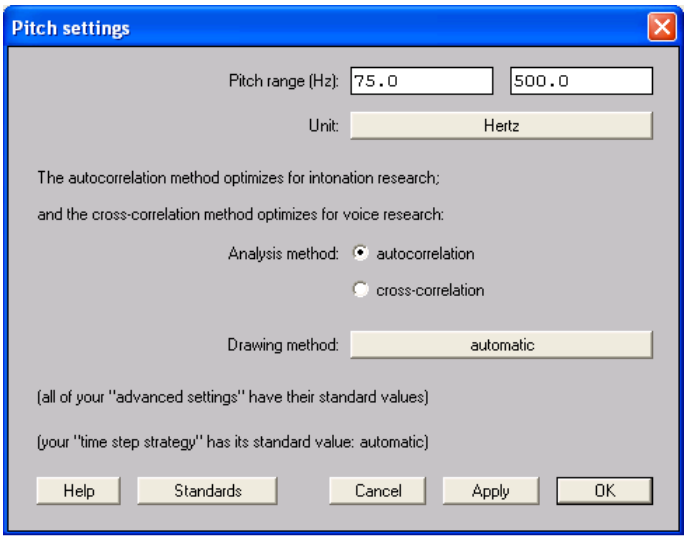


Figure 2. *Pitch Settings Interface.*

The settings in this window are only basic parameters, in the advanced pitch settings it is possible to modify the maximum number of candidates, the silence threshold, the voice threshold, the voiced/unvoiced cost etc. All these parameters are fairly complicated and require a good level of acquaintance with Praat (Halliday and Greaves 2008: 18-36).

In pitch analysis, the variation of pitches is particularly interesting because it gives an accurate and precise description of the fall and rise in the voice. The incidence of rising or falling tones can be measured by indicating how sharply the rise or the fall occurs. The indication of the gradual or abrupt modality of change in the sound allow us to gain considerable information. For this reason, in the analysis of utterances, it is fundamental to consider the highest and lowest pitch in every segment, just to give a clear idea of the movement. The waves in the spectrograms are a visual representation of the swings and oscillation of sounds that are clearly described by pitch analysis and variations (Halliday and Greaves 2008: 18-36).

2.3. Duration

Duration has to do with the much familiar parameter of time. In Praat, time is measured in milliseconds and it is represented on the horizontal X axis. To measure the duration of utterances is crucial because it permits to compare how long it takes a speaker to pronounce a sentence.

Each utterance should be divided into segments in order to measure the duration of each individual segment and the duration of the whole utterance. Generally speaking, we presume that NSs pronounce an utterance faster than NNSs for different reasons. First, it is obvious that NNSs speak slowly because they are less confident than NSs; second, NSs have the tendency to shorten words and to contract some sounds into a single one; third, speech speed is a also a personal trait since there are people who speak faster than others. For all this reason, NNSs are expected to speak slower than NSs.

3. ToBI: Tone and Break Indeces

ToBI is one of the most popular systems of annotation of intonation[3]. It was originally created by Pierrehumbert (1987) but it has been re-elaborated and modified in later studies (Pierrehumbert and Hirschberg, 1990; Silverman et al. 1992; Syrdal et al. 2001). According to Halliday and Greaves (2008: 12), it has been 'proposed by Silverman et al. (1992) as an agreed system for transcribing prosodic structures which could be used consistently by researchers in various fields'. This model of intonation description is used to talk about prosodic phenomena, 'to allow researchers to compare their findings more easily, within and across languages, and to facilitate the construction of very large speech corpora, especially for learning associations between prosodic features and other aspects of texts' (Hirschberg 2002: 31).

Beckman and Elam (1997) explain that ToBI consists of four level of analysis: an orthographic tier, a break index tier, a tonal tier and a miscellaneous tier (see also Cruttenden 1997: 59 and Hirschberg 2002: 33, 2004: 2).

The center of the linguistic analysis is especially the tonal tier, 'where pitch accents, phrase accent, and boundary tones describing targets in the fundamental frequency (f0) define intonational phrases' (Hirschberg, 2002: 33). So, the tonal tier is mainly concerned with the pitch movements of pitch accents, phrase accents and boundary tones. These movements or pitch tendencies are identified as either H (high) or L (low). So ToBI 'rather than analyzing intonation patterns in terms of pitch contours (rise, fall, fall-rise, etc.), [it] breaks them down into components, basically High and Low in various combinations' (Wells 2006: 261).

The H and L marks are associated to pitch accents, phrase accents and boundary tones; all these elements determinate the intonation contour. 'The meanings of intonation contours are said to be compositional in the sense that each tone in any sequence contributes separately to the overall meaning and the meaning of the whole is equal to the sum of the parts' (Cruttenden, 1997: 64). Therefore, sequences of H and L marks describe the specific contours of intonation patterns by providing compositional meaning.

3.1. Interpreting tones

The intonation of a sentence is clearly determined by the clause structure but also by the speaker's purpose. Due to the fact that intonation patterns are not predictable, the interpretation of tones is a difficult task to achieve.

In the following table, it is evident how the same word (e.g. *Amelia*) can assume a different tonal tier according to its speech function in the discourse.

[3]A comprehensive description of the ToBI system can be found in the ToBI conventions document and the training materials available at http://ling.ohio-state.edu/tobi, (written by Beckman and Elam 1997). This model is based on earlier work by Pierrehumbert (1987) and Pierrehumbert and Hirschberg (1990).

Table 1. *ToBI transcription of samples utterances.* (from Ladefoged, 2006: 125).

1)	A′melia.	TONE TIER	[H* L-L%]
	Simple statement in response to *What is her name?*	SEGMENTAL TIER	[ə m i: l i: ə]
2)	A′melia?	TONE TIER	[L* H-H%]
	A question equivalent to *Did you say Amelia?*	SEGMENTAL TIER	[ə m i: l i: ə]
3)	A′melia ─	TONE TIER	[L* L-H%]
	Addressing Amelia, indicating that it is he turn to speak.	SEGMENTAL TIER	[ə m i: l i: ə]
4)	A′melia!?	TONE TIER	[L+H* L-H%]
	A question indicating surprise.	SEGMENTAL TIER	[ə m i: l i: ə]
5)	A′melia!!	TONE TIER	[L+H* L-L%]
	A strong reaction, reprimanding Amela.	SEGMENTAL TIER	[ə m i: l i: ə]

As it is clear from the table, while the segmental tier is unvaried, the tone tiers change according to the contextual meaning of the word. In (1), *Amelia* represents the answer to a question and it has a standard declarative tone. In (2), since the utterance produced is equivalent to the question *Did you say Amelia?*, the tone tier is that of standard 'yes-no' questions. The difference between (3) and (4) is that in (3) *Amelia* is pronounced with an encouraging attitude while in (4) it conveys a surprised tone. The tone people assume when they address someone as to encourage him or her to speak is different from the tone used to express surprise or astonishment. In the first case the tonal tier represents a low rise, in the second case it represents a high rise. In (5), *Amelia* is pronounced with an irritated tone and it describes a strong reaction. For this reason, the tonal tier is based on a sharp rise and a low fall.

4. The model by Ladefoged

In the linguistic analysis provided by Ladefoged (2006), the H and L labels describe four levels of pitch movements: optional pre-nuclear pitch accents, nuclear pitch accents, phrase accents, and boundary tones. So, compared to the system of Pierrehumbert (1987, 1990) and Hirschberg (1990, 2002), Ladefoged introduces a new optional element: the optional pre-nuclear pitch accent. The ToBI system for characterizing English intonation has been elaborated by Ladefoged as follows:

Table 2. *The ToBI system for characterizing English intonation.* 'Each intonational phrase (tone group) must have one item for each of the last three columns, and may also have additional pitch accents marked on other stressed syllables, as shown in the first column'. (from Ladefoged, 2006: 225)

Optional pre-nuclear Pitch Accents on Stressed Syllables	Nuclear Pitch Accent	Phrase Accent	Boundary tones
H*	H*		
L*	L*		
L + H*	L + H*	L -	H %
L* + H	L* + H		
H +!H*	H +!H*	H -	L %
(!H*)	(!H*)		

H and L respectively represent high or low pitch. H* (read 'H star') and L* (read 'L star') are pitch accents which are typically written on a line called tier (Ladefoged 2006: 125). As for the asterisk, it indicates the central part of the pitch accent and it is never positioned on a phrase accent or on a boundary tone. 'The phrase accent is the component between the last pitch accent and the boundary tone and it is represented by H- or L- without any diacritic' (Halliday and Greaves 2008: 12). In order to differentiate boundary tones from other kinds of pitch and phrase accents, boundary tones are indicated with a percentage sign (%). This symbol identifies the boundary tone as the last component of a tone. The boundary tone is marked as H% or L% depending on whether the utterance ends with a rising or falling pitch (Ladefoged 2006: 125).

Generally speaking, a standard declarative intonation pattern ends in a low phrase accent and a low boundary tone (L-L%); a standard 'yes-no' question ends in a high phrase accent and a low boundary tone (H-H%) (Hirschberg 2004: 2 and Ladefoged 2006: 127). Even though it is difficult to generalize, we can assign to every phonetic variants an usual target. For example, the single tones H* and L* indicate a peak accent or a low accent. The combined tones L*+H and L+H* designate respectively a scooped accent and a rising peak accent. As for sentence stress, the phrase accent L- and H- suggest the tonal development of the sentence. The low sentence final is indicated by L%, while H% means that 'the boundary tone is high at the end of a constituent' (Martin 2004: 2).

5. Comparison between ToBI and the British nuclear tone framework

The ToBI system and the British nuclear tone framework are clearly different because the first one describes the intonation patterns by using *H* and *L* annotation, while the second one explains the development of the sentence intonation by *rise* and *fall* annotation. Nevertheless, as Toivanen (2005:2) argues, some examples of correspondences between the two systems have been noticed and can be 'represented as follows: high-fall (H* L-L%); low-fall (L* L-L%); high-rise (L* H-H%, H* H-H%); low-rise (L* H-L%); level tone (H* H-L%)'.

The British nuclear tone framework is based on the assumption that information is provided by spoken material divided into chunks. 'These chunks are known as intonation phrases or IPs. Each IP in an utterance has its own intonation pattern' (Wells 2006: 6). So, the specific tone associated to each IP determines the representation of the contours. Obviously, if an IP performs an abrupt rise, the contour will be a high rise; if the IP performs a fall followed by a rise, the contour will be a fall-rise, etc. Notably, the typical IP can be composed by four different elements: the pre-head, the head, the nucleus, and the tail. The tail is, by definition, the last part of the IP and contains no stressed syllables. Also the pre-head, the first part of the IP, is never stressed. By contrast, both the head and the nucleus carry stress. The head is the most prominent stressed word while the nucleus is in the location of the last stress in the word (Wells 2006: 7-9).

The British nuclear tone system has been outclassed by the ToBI system because ToBI makes the results of speech synthesis easier to be compared. It is still difficult to decide which system is superior to the other because both systems present positive and negative aspects. Some of the main points of comparisons are summarized by Cruttenden (1997: 64) who asserts that 'ToBI captures some sentences […] better than nuclear analysis; in particular it captures the relationship between level and non-level contours'. In addition, Cruttenden argues (1997: 66) that 'the transcription of text using H's and L's is an altogether much more difficult affair than using the tone marks most widely used in nuclear tone analysis.' So, on the one hand ToBI results to provide a better analysis of the intonation patterns; on the other

hand, this system is more complex than the British nuclear tone system.

ToBI and the British nuclear tone framework are undoubtedly different. However, despite appearances, there is much in ToBI which is comparable to the earlier traditional British nuclear system. By analyzing the table below, elaborate by Ladd (1996:82), it is possible to find some correspondences:

Table. 3. *Correspondences between Pierrehumbert annotation system and British-style nuclear tone.* (from Ladd, 1996: 82).

Pierrehumbert		**British-style**
H*	L-L%	Fall
H*	L-H%	Fall-rise
H*	H-L%	Stylized high rise
H*	H-H%	High rise
L*	L-L%	Low fall
L*	L-H%	Low rise (narrow pitch range)
L*	H-L%	Stylized low rise
L*	H-H%	Low rise
L+H*	L-L%	Rise-fall
L+H*	L-H%	Rise-fall-rise
L+H*	H-L%	Stylized high rise (with low head)
L+H*	H-H%	High rise (with low head)
L*+H	L-L%	Rise-fall (emphatic)
L*+H	L-H%	Rise-fall-rise (emphatic)
L*+H	H-L%	Stylized low rise
L*+H	H-H%	Low rise
H+L*	L-L%	Low fall (with high head)
H+L*	L-H%	Low fall-rise (with high head)
H+L*	H-L%	Stylized high rise (with high head)
H+L*	H-H%	Low rise (high range)
H*+L	H-L%	Stylized fall-rise
H*+L	H-H%	Fall-rise (high range)

The table above shows twenty-two combinations between the system elaborated by Pierrehumbert in 1980 (an earlier version of ToBI) and the British nuclear tone framework. Despite the distinctions between the two systems, it is clear that there are some correspondences between the two systems. However, there is not a perfect parallelism because the two systems are based on different descriptive criteria. As Ladd points out (1996: 83), 'the grouping based on the Pierrehumbert analysis shows five completely parallel sets of four types, plus two additional ones, whereas from the point of view of the British tradition certain types like 'low rise' and 'high rise' show up rather unpredictability at several different places in the table, and references to pitch range or to the preceding head are required here and there to describe certain distinctions'.

While ToBI has been specifically tailored to the analysis of American English, the British nuclear tone system has been modeled on the intonation features of British English. However, in recent years, ToBI has been adapted also to other languages. It 'has been "localized" for languages such as German, Japanese, Korean, Greek and there are ToBI systems under development for at least Serbo-Croatian, Mandarin, Cantonese, and Spanish' (Toivanen 2005: 1). Actually, researchers are currently working for an application of ToBI to Italian.[4]

[4]Grice, Savino, D'Imperio and Avesani (2005) conducted a study aimed at finding some strategies in order to create a variety of ToBI modeled on the Italian intonation system. In particular, they studied specific strategies

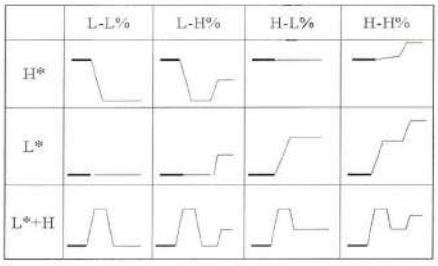
5.1. Drawbacks of the ToBI approach

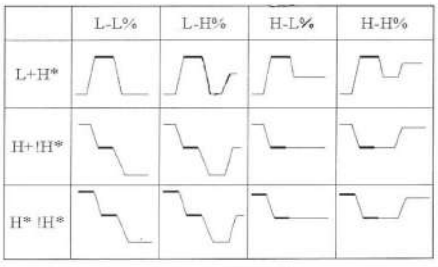
The ToBI system was devised especially to annotate speech corpora and to work in speech technology. Even though it 'continues to provide one of the most fundamental links between speech and language technology research on the one hand and basic research in phonetics and linguistic on the other' (Bird and Harrington 2001: 4), the effectiveness and accuracy of ToBI are still discussed and are nowadays an open question. The deficiencies of ToBI that have been pointed out (Pitrelli et al. 1994; Bird and Harrington 2001; Martin 2004; Toivanen 2005) have to do with the fact that

> there are three main parts to consider when transcribing intonation: dividing an utterance into one or more prosodic phrases; deciding which word is the nuclear accented word and which of the remaining words in the utterance are accented or unaccented; and finally assigning a tune, consisting of one or more pitch accents and a boundary tone to each prosodic phrase' (Harrington 2008: 1).

These three aspects of intonation transcription are crucial for the identification of the correct intonation patterns. Martin (2004: 4) blames ToBI because 'it mixes phonetic and phonology' in the analysis. Moreover, Martin (2004: 4) demonstrates that a series of melodic variations are annotated with the same transcription even though the amplitude and the duration of melodic variations are different in each case. Despite all this criticism, Martin himself confirms that ToBI is 'extremely popular, to the point that prosodic studies not using it are often neglected and discarded from the research mainstream by the research community' (2004: 3). So, this annotation system is, anyway, the prevailing model of prosodic annotation broadly accepted and largely favored.

Table 4. *ToBI Contours for Standard American English.* Schematic representation of all the possible combinations which can occur in Standard American English (elaborated by Hirschberg, 2002).





In order to overcome the difficulties related to the process of annotation and identification of patterns, Hirschberg elaborated a

for ToBI labelling varieties of Italian. In addition, Grice (1995) carried out some analysis about the ToBI system applied to the description of the intonation of interrogation in Palermo Italian.

schematic representation of the possible contours in the ToBI system (see table 4). This representation is a visual survey of all the possible combinations which can occur in Standard American English.

The series of tone contours are stylizations of the curves which appear in spectrograms elaborated with Praat. A combined use of a software for computerized signal analysis, such as Praat, and an annotation system, such as ToBI, have enabled the accurate study and analysis of intonation and prosodic structures of spoken utterances. For this reason, despite its deficiencies, ToBI is by far one of the best annotation systems available because of the complete correspondence of graphic representation to phonetic reality and to semantics and pragmatics (Cruttenden 1997: 64).

6. Conclusion

The development of speech analysis software and the creation of hundreds of annotated speech corpora have many positive effects not only in the field of linguistic research but also in language learning (especially in second language acquisition). The number of collections of materials available and the different approaches discussed in a series of international workshops testify the real need to advance these studies for their theoretical and applied implications.

Overall, speech analysis software are raising a new awareness on the importance of these tools in language learning and linguistic research. Linguistics, computational linguistics and speech engineers have increasingly looked at technology as 'an awareness-raising tool for sensitizing teachers to features of stress and rhythm in English' (Coniam 2002: 30). Speech analysis software have undoubtedly potentialities that can be exploited in the domain of education.

Since the importance of collecting, recognizing and analyzing prosodic information is ever more needed in this research domain, it is recommend that more extensive research is done. The need to implement speech analysis software as well as to conduct exhaustive studies on annotation systems is of primary importance to improve communicative competence in L2 speakers.

References

Anderson-Hsieh, J. 1994. Interpreting Visual Feedback on Suprasegmentals in Computer Assisted Pronunciation Instruction. *CALICO Journal*, vol.11, 4.

Beckman, M. and E. Ayers. G. 1997. Guidelines for ToBI labeling. The Ohio State University. Available at: http://ling.ohiostate.edu/Phonetics/EToBI/homepage. html. Accessed 22 November 2010.

Bird, S. and J. Harrington. 2001. Speech Annotation and Corpus tools. Speech Communication. Available at: http://www.phonetik.uni-muenchen.de/%7Ejmh/ research/ papers/ bird01.specom.pdf. Accessed 22 November 2010.

Boersma, P. and D. Weenink. 2001. Praat: a System for Doing Phonetics by Computer. Tech. Report 132, Inst. Phonetic Sciences, Univ. Amsterdam. Available at http://www. fon.hum.uva.nl/praat/. Accessed 22 November 2010.

Busà, M. G. 2008. Teaching Prosody to Italian Learners of English: Working towards a New Approach. *Ecolingua. The Role of E-corpora in Translation and Language Learning*, 113-126.

Chun, D. M. 2002. *Discourse Intonation in L2: From Theory and Research to Practice*. Amsterdam: John Benjamins.

Chun, D. M. 1998. Signal Analysis Software for Teaching Discourse Intonation. *Language Learning and Technology* 01, Vol. 2, 61-67.

Coniam, D. 2002. Technology as an Awareness-raising Tool for Sensitising Teachers to Features of Stress and Rhythm in English. *Language Awareness* 01, Vol. 11, 30-42.

Cruttenden, A.1997. *Intonation*. Cambridge: CUP.

Greaves , W. 2008. Intonation in Systemic Functional Linguistics. I n H. Ruqaiya, C. Matthiessen, and J. Webster (eds.). *Continuing discourse on language : a functional perspective* . London: Equinox.

Grice, M. 1995. *The Intonation of Interrogation in Palermo Italian: Implications for Intonation Theory.* Tübingen: Niemeyer.

Grice, M., M. P. D'Imperio, M. Savino and C. Avesani. 2005. Towards a Strategy for ToBI Labelling Varieties of Italian. In J. Sun-Ah (ed.). *Prosodic Typology: The Phonology of Intonation and Phrasing* Oxford: Oxford University Press.

Halliday, M.A.K. and W. S. Greaves. 2008. *Intonation in the Grammar of English*. London: Equinox.

Harrington, J., S. Cassidy, T. John and M. Scheffers, 2003. Building an Interface between EMU and Praat: a Modular Approach to Speech Database Analysis. International Congress of Phonetic Sciences. Available at: http://www.phonetik.uni-muenchen. de/~jmh /research/ papers/PS022037.pdf. Accessed 22 November 2010.

Harrington, J. 2008. Intonation: Transcribing with ToBI. Sidney. Macquarie University. Available at:http://clas.mq.edu.au/speech/phonetics/phonology/intonation/tobi_transcribe. html. Accessed 22 November 2010.

Hirschberg, J. 2002. Communication and Prosody: Functional Aspects of Prosody. *Speech Communication* 36, 31-43.

Hirschberg, J. 2004. Pragmatics and Intonation. In L.R. Horn, and G. Ward (eds.). *The Handbook of Pragmatics*. Oxford: Basil Blackwell, 515-537.

Ladd, D.R. 1996. *Intonational Phonology*. Cambridge: CUP

Ladefoged, P. 2003. *Phonetics Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Oxford: Blackwell.

Ladefoged, P. 2006. *A course in Phonetics*. Boston: Thomson.

Martin, P. 2004. Intonation and Syntax: Another Point of View. Available at: http://www. chass.utoronto.ca/epc/srb/cyber/martin1.pdf. Accessed 22 November 2010.

Pierrehumbert, J. 1987. *The Phonology and Phonetics of English Intonation.* (Ph.D thesis, 1980). Indiana University Linguistics Club.

Pierrehumbert, J. and J. Hirschberg. 1990. The Meaning of Intonational Contours in the Interpretation of Discourse. In P. R. Cohen, J. Morgan and M. E. Pollack (eds.), *Plans and intentions in communication and discourse*. Cambridge, MA: MIT Press, 271-311.

Pitrelli, J. M. Beckman, J. Hirschberg. 1994. Evaluation of Prosodic Transcription Labelling Reliability in the ToBI Framework. *Proceedings of the 1992 International Conference o Spoken Language Processing*, 123-126.

Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert, and J. Hirschberg. 1992. ToBI: a Standard for Labeling English Prosody. *Proceedings of the 1992 International Conference o Spoken Language Processing*.

Syrdal, A.K., J. Hirschberg, J. McGory and M. Beckman 2001. Automatic ToBI Prediction and Alignment to Speed Manual Labeling of Prosody. *Speech Communication* 33, 135-151.

Toivanen, J. 2005. ToBI or not ToBI? Testing Two Models in Teaching English Intonation to Finns. University of Oulu and Academy of Finland. Available at: http://

www.phon.ucl.ac.uk/ home/johnm/ptlc2005/pdf/ptlcp29.pdf. Accessed 22 November 2010
Wells, J.C. 2006. *English Intonation. An introduction*. Cambridge: CUP.

ABSTRACT

Instruments and methods for the analysis of prosody

Key words: speech analysis software, annotation system, CALL, speech corpus.

The development of speech analysis software and the creation of a huge number of annotated speech corpora have positive effects not only in the field of linguistic research but also in language learning (especially in second language acquisition). Speech analysis software and speech annotation systems allow researchers to conduct experiments such as contrastive analyses of intonation patterns produced by native speakers (henceforth NSs) and non-native speakers (henceforth NNSs). Praat and ToBI are useful also because they can be used in students' instruction: the visualization of sound waves, spectrograms, pitch lines and their annotations provide a feedback that helps students and researchers to analyze and evaluate speech productions.