1. Introduction: English language teaching in a global world

The expanding role of English as the language of international communication has important implications for English language pedagogy. Statistics show that speakers of English as a second and foreign language worldwide largely outnumber the speakers of English as a first language, so that International English can no longer be considered the ‘exclusive property’ of its native speakers (Ferguson, 2006; Kachru, 1985; Trudgill, 1998). At the theoretical level, the need is felt for revisiting traditional definitions of standards, norms and models to use in language descriptions and teaching, so as to integrate varieties of English spoken by non-native speakers, and developed both within well-established traditions, e.g., Indian English, and new traditions, e.g. European English (also called EuroEnglish, see Jenkins, 2006; Seidlhofer, 2004). At the applied level, new teaching practices are emerging, which represent “a ‘paradigm shift’ away from conventional EFL models” (Graddol, 2006: 15), and emphasize skills and abilities that can meet the evolving needs of International English users. For example, because of the large number of English non-native speakers worldwide, the view is being accredited that learners should be exposed to different varieties of English, including non-native ones.

The current debate on English as an International Language is placing pronunciation in a new position with regard to language instruction. This is due to the fact that interlanguage communication rests on the concept of mutual intelligi-
bility, and pronunciation is one of the main factors contributing to it. In other words, instruction should emphasize pronunciation to ensure that L2 learners’ accent in English does not interfere with their ability to make themselves understood in interactions with speakers from different linguistic backgrounds. To this purpose, studies are being carried out internationally to investigate characteristics and dynamics of interlanguage communication in relation to foreign accent and intelligibility (Jenkins, 2000; Seidlhofer, 2004; Munro and Derwing, 1995a; 1995b; Munro et al., 2006; Patil, 2006; Pickering, 2006).

This paper discusses aspects of Italian pronunciation in English which may affect intelligibility, and reviews a method for the teaching of prosody which was implemented experimentally at the University of Padova, showing great potential for improving in-class and at-home pronunciation practice.

2. Trends in pronunciation teaching

Attitudes and approaches to English pronunciation teaching have changed considerably over the past fifty years. Since the advent of the communicative approach, pronunciation has been recognized a key role in improving the learner’s oral skills, and contributing to the success of oral communication (Anderson-Hsieh, 1989; Celce-Murcia, 1987). In fact, by improving the learner’s oral skills, an accurate pronunciation is believed to help the learner increase self-confidence, promote social interactions outside the classroom (Cunningham Florez, 1998; Morley, 1991), and contribute to clarity and efficiency in professional exchanges (Neri, Cucchiarini, Strik and Boves, 2002). Also, because non-native pronunciation may be socially stigmatized and contribute to the negative stereotyping of some second-language learners, pronunciation teaching is viewed as a means to improve the learner’s personal social acceptance, decreasing the odds of social or professional discrimination (Derwing, Rossiter and Munro, 2002; Munro, 2008).

In the last decade, with the change of the role of English in the world, the objectives of pronunciation teaching have changed too. Traditionally, the aim of pronunciation teaching was the eradication of any trace of ‘foreign accent’ and the model varieties of English to imitate were British or American English. Today, in a world characterized by an infinite number of English accents, promoting a ‘perfect English accent’ appears unsustainable. Rather, the emphasis has shifted on pronunciation that is intelligible, and the focus on instruction is a list of features which are deemed particularly detrimental to intelligibility if non-corrected (e.g., Brown, 1989; Derwing and Munro, 2005; Jenkins, 2000; Richards and Renandya, 2002; Tarone, 2005).

However, the notion of intelligibility in L2 speech is far from being well defined (Pickering, 2006), as it is highly dependent on non-linguistic factors such as degree of exposure or familiarity with a certain variety of English, listeners’ attitudes towards second language speakers, etc. Research has shown that L2 learners are often unable to produce the characteristics of all individual L2 sounds and their allophonic variations, and this leads to the perception of ‘foreign accent’. But it is not clear whether it is the segmental vs. suprasegmental aspects of L2 speech which are more likely to affect L2 speakers’ intelligibility.
In fact, studies in second language phonology have lagged behind studies of other aspects of learners’ language, such as morphology, syntax, discourse or pragmatics (Tarone, 2005), and research has particularly neglected the phonology of second language suprasegmentals, though there is an increasing awareness that they play an important role in second-language acquisition processing, and in the perception and judgment of L2 speech (for a review, see Trofimovich and Baker, 2006). More systematic investigations are needed to examine how L2 suprasegmentals are learned, what factors influence their learning, and how they affect classroom-based and/or naturalistic interactions. This kind of research will provide essential data for language teachers to improve their teaching methods and materials, and to enhance learners’ communication skills.

3. Italian and English phonological systems: a comparison

A good starting point for teaching pronunciation is to single out L2 learners’ most recurrent problems and work from there.

Notoriously, one of the big problems for Italian learners of English is the production of vowels. This shows up in both the Italians’ inability to differentiate words distinguished in English solely by the vowel (e.g., sheep – ship, beg – bag, etc.), which may lead to intelligibility problems, and their tendency to have a little epenthetic vowel at the end of English words, especially those ending with a stop consonant (e.g., in did, big, etc.), a feature which is often emphasized in stereotypes of the Italian accent in English. In fact, it has been shown that the production of English vowels by Italians largely correlates with Italian speakers’ perceived degree of accent in English (Busà, 1995; Flege et al. 1999; Flege et al. 2003; MacKay et al. 2001; Piske et al. 2002). It would seem obvious, then, that teaching English pronunciation to Italian native speakers should start from the English vowels.

But what exactly is involved in the production of English vowels, and how can Italians be taught to produce vowel differences not existing in Italian?

Italian speakers’ difficulties in producing English vowels stem from differences existing between the Italian and English phonological systems, both at the segmental and suprasegmental levels. At the segmental level, English has 11-13 vowels in its inventory, depending on the variety of English under consideration, while Italian has only seven vowels. The fewer vowels in the Italian system condition the Italian speakers’ production and perception of English vowels, and lead to frequent hypo-differentiations of vowel contrasts. In addition, in English vowels may span from full to reduced, in both quality and duration, and even disappear, depending on the degree of stress they receive in the utterance; in Italian, vowel quality tends to remain quite stable, regardless of the degree of stress on the vowel or any other phonological condition of the utterance. Thus, in English, phonological rules operating at the level of suprasegmentals (i.e., syllable structures, rhythmic tendencies, stress assignment rules, and intonation) trigger vowel reduction processes and create distinctions between vowels in ‘strong’ and ‘weak’ syllables. In Italian, these rules do not operate: syllables tend to have the same ‘weight’, and vowels are always fully pronounced.
The pronunciation of English vowels is not the only major problem Italian learners of English are faced with. Preliminary studies (Busà, 2007 and forthcoming) suggest that Italian learners of English may be unable to convey the appropriate English prosodic information, with a possible effect on the outcomes of their communication in English. That inappropriate prosody may affect overall communication is not surprising, given the fact that prosodic features have been found to play a major role in the production and perception of foreign accent (Boula de Mareüil and Vieru-Dimulescu, 2006; Munro and Derwing, 1995a). Also, prosody is a very important part of speech, as it conveys linguistic and pragmatic meaning. In both Italian and English, though with different modalities, prosody is used in the disambiguation of structurally ambiguous sentences, to signal the information status of an utterance (i.e., given vs. new information, emphasis or contrast, etc.), and to define speech functions (for example by differentiating between statements, questions, requests, etc.). Prosody may also convey paralinguistic information, for example with regard to the emotional state of the speaker (e.g., anger, happiness, love, etc.), the truth value of the proposition (e.g., certainty vs. uncertainty) or the level of the speakers’ engagement (i.e., when the speaker is seeking support, responding to something, anticipating possible responses and objections, etc.).

A comparison of English and Italian prosodic features reveals what could be the possible pronunciation issues for an Italian speaker of English. The major differences in the Italian and English suprasegmental features are summarized below.

At the syllabic level, English has mainly CVC-type syllable structures, and allows complex consonant clusters both in syllable initial and final position; Italian has mainly CV-type syllable structures, with a distribution of long vowels in open syllables and short vowels in closed syllables; it does not allow complex consonant groups in syllable-initial or final position, and allows only a limited set of consonants in word-final position. As far as rhythm is concerned, English has been referred to as a stress-timed language, and Italian as a syllable-timed language (for a review, see Busà, 1995). This means that English will show a tendency to keep intervals between stresses equal, independently of the number of intervening unstressed syllables, by compressing sequences of unstressed syllables; Italian, on the other hand, will show a tendency to keep syllables at about constant duration, and more syllables (stressed or unstressed) will proportionally increase the duration of the sentence (Schlüter, 2005). Thus, whereas English is characterized by full vowels in stressed position and (highly) reduced vowels in unstressed position, Italian has no vowel reduction at the phonological level and limited vowel reduction at the phonetic level (Farnetani and Busà, 1999).

English and Italian also differ markedly in the way intonation is used to signal discourse information structure and focus, as well as in the intonation patterns used linguistically. In the first place, this difference concerns the relation between word order and intonation. Word order and intonation are the two most commonly used focus marking devices, and languages differ in the preference for one over the other, and in the ways in which the position of the focal constituent in the sentence may affect the intonational realization of focus (Chen et al. 2007). English has few inflections and a relatively fixed word order, and it relies heavily
on intonation to convey grammatical information or focus elements in the sentence. Italian, on the other hand, has more inflections and a more flexible word order than English, and so provides its speakers with the option of giving prominence to some information by rearranging words in the sentence. In addition, English uses intonational accent (or extra stress) to mark grammatically salient elements (for example new or emphatic information) as prominent, while given or old information is de-accented; typically, focus accent in English is found on the last major word of the sentence, but can come earlier to emphasize one of the earlier words or to contrast it with something else. In Italian, prosody is not used to distinguish between new and given information, that is, givenness is not prosodically marked by deaccenting elements carrying given information or by using a particular type of pitch accent; rather, prominence is given to elements that are in focus (Avesani and Vayra, 2005; Bocci and Avesani, 2008).

Thus, following the strategies of their native language, Italian speakers of English will be unable to mark salient discourse information through intonation, and show instead a tendency to either move syntactic elements around in the sentence, or use other linguistic devices (for example lexical items) to mark discourse focus.

4. Teaching pronunciation the traditional way

The traditional methods for teaching pronunciation to L2 learners treat segments and suprasegmentals as separate component parts of pronunciation. Lessons focusing on segments typically involve simplified explanations of the articulation of individual sounds, as well as refinement of the learners’ perception of the new sounds through aural discrimination exercises and minimal pair drills, that is, exercises where the learner is presented with the sounds in contrast. However, different sounds pose different types of difficulties for learners, though they are often treated in the same way in many pronunciations books. For example, articulatory diagrams are used to explain the position of the articulators during the production of a sound. Such diagrams may be a source of accurate information for the production of consonants, which require an obstruction or approximation of the articulators in specific places of the vocal tract, but they certainly do not provide much help for the production of vowels, which require the mobile articulators only to approach the fixed articulators. Unlike consonants, vowels are hard to teach because learners cannot ‘feel’ how to position the tongue correctly to create vowel distinctions, and instruction can only describe vowel production in relative terms. Learning the L2 vowels rests largely on perception. In this sense, minimal pair drills are highly effective in raising the learners’ awareness of the differences between the L1 and the L2 vowel phonemic inventories, while diagrams are more conducive to favoring an enhancement of the learners’ production for consonant sounds.

In general, even when learners do come to some understanding of the segmental differences existing between the native language and the L2, producing them accurately in their spoken language is quite a different story. In fact, in order for perception exercises (like minimal pair drills) to be effective for im-
proving the learners’ articulation of new sounds, learners should get individual feedback on their productions, as their self-correcting ability is limited by the influence of their native language (Carey, 2005). But this is a highly unlikely situation in classrooms with only one language teacher. Generally, rather than developing new features, learners will tend to revert to the features which are found in their native language. For example, Italians who are trying to create a distinction between pairs like ‘bead’ and ‘bid’ may either end up exaggerating the duration of the vowel in ‘bead’ to produce oppositions like [biːd] and [bid], or, in the attempt to produce a distinction between the tense and lax vowels in the target words, revert to the closest contrast existing in Italian, that is [i] and [e], thus produce vowels like the one in ‘bid’ as [bed], or sometimes even [bed].

Another likely outcome of pronunciation exercises focusing on segmentals is hypercorrection mechanisms, that is the over articulation of the target sounds. These may occur at the expenses of other processes which are important for successful communication in English. For example, a problem arising when the production of full English vowels is emphasized is that, anxious to produce vowel distinctions not existing in their native language, learners may pronounce all vowels in their full, citation forms, rather than aiming at differentiating between full and reduced vowels. This contributes to increasing a well-know difficulty (i.e., vowel reduction processes) for learners of syllable-timed languages like Italians, who are particularly prone to pronounce all English vowels as full vowels.

It is unquestionable that a knowledge of L2 word stress and intonation patterns will help English learners to pronounce both full and reduced vowels, as well as assign a more stress-timed rhythm to sentences, giving prominence to semantically loaded words. As for sound perception, knowing that English vowels may present themselves in highly variable forms -- from full to reduced, by effect of word and sentence stress and intonation, will help learners to associate actual spoken words to their citation forms, and thus will help learners understand the spoken language.

Thus, because, in speech, segmentals and suprasegmentals overlap and contribute to each other in many important ways, in pronunciation classes they should be taught together rather than separately. Focusing on stress, rhythm and intonation can help learners to improve their overall pronunciation, and to sound more natural, and can lead to more comprehensible speech as well as better understanding of other people’s speech. In fact, it has been found (Derwing, Munro and Wieber, 1998) that learners are more likely to learn to produce segmentals effectively in spontaneous speech if they had instruction emphasizing suprasegmental features than if they only received instruction on the L2 segmentals (i.e., vowels and consonants).

However, teachers report a general difficulty teaching some features of prosody, mostly because they feel they lack the amount of competence necessary to embark in a such a cumbersome task as is required for describing all the variations in pitch, stress and rhythm part of the English prosodic system. As with all aspects of pronunciation, teaching L2 suprasegmentals requires some knowledge of the L1 and the L2 prosodic systems. In addition, to make teaching more fruitful, it is important to distinguish between “what English speakers do” and
“what learners of English need to learn” (Roach, 1996: 47), that is what L2 features should be taught and corrected and which do not need to be because they are not relevant for effective communication. This requires a good knowledge of how intonation in discourse functions in the L1 and L2.

Unfortunately, no systematic contrastive studies of Italian and English prosody have been carried out, and we are largely unaware of how Italians transfer their prosodic patterns into their L2 and how this affects their communication in English. Currently, some studies are being conducted at the University of Padova, which seem to be yielding promising results. The following section illustrates how the method used can both give indications of pronunciation problems for the learners of English and be used effectively in language pronunciation classes.

5. TOWARDS A NEW METHOD FOR TEACHING PROSODY TO ITALIAN LEARNERS OF ENGLISH

As seen in section 4, one of the problems with many pronunciation drills in the language classroom is that learners have no precise feedback they can rely on to compare and contrast their productions in L2 with the native speakers’ target productions. However, improvement in the L2 is strongly connected to the amount of feedback learners can get, since influences from L1 prevent them from discriminating or reproducing L2 sounds and phonological processes. In addition, as regards intonation properly, it appears difficult for learners (and for native speakers too) to gain a ‘conscious awareness’ of pitch movements, both in production and perception. This may affect learners’ ability to hear and produce native-like intonation, though they may have no difficulty producing appropriate pitch fall or rise in statements and questions, since all languages use falling and raising pitch linguistically (Chela-Flores, 2003).

To provide learners with pronunciation feedback, today’s tendency in pronunciation pedagogy is to expose learners to a wide variety of techniques, also involving multisensory models of presentation (e.g., visual, auditory, tactile, physical and affective senses, see Underhill, 2005). Some of these techniques are aimed at letting learners compare the physical characteristics of the L2 target with the closest native L1 phoneme, so that learners can get a real measure of the distance of their own productions from the target phones. By allowing learners to acquire a conscious awareness of the differences between the L1 and the L2, these techniques stimulate improvement through self-monitoring and self-correction.

A technique that is becoming increasingly popular among pronunciation teachers is the use of speech analysis software in L2 pronunciation classes as a source of audio-visual feedback for students’ productions. The use of speech analysis software allows learners to record and visualize their speech output on their computer monitors to obtain real-time information about the acoustic properties of this output. These visualizations can be used by both learners and teachers to compare and evaluate learners’ productions with those of native speakers, and learners can get an awareness of what details of their productions they need to change to approach the native speakers target. In addition, by recording and visualizing their pronunciation over time, learners can get a concrete idea of their
progress. This method has been considered highly effective, and particularly for speech prosody. Through it, learners can easily learn about L1 and L2 segments durations, rhythmic tendencies, reduction processes, pitch and intonation contours, by learning to associate the patterns on the display with the actual sounds, having only minimal knowledge of acoustic phonetics (e.g., Carey, 2005; M. Chun, 1998; De Bot, 1983; Eskenazi, 1999; Lambacher, 1996a, 1996b; Spaai and Hermes, 1993; Stibbard, 1996; Wennerstrom, 2000).

This section will show an example of how this method has been used and what results can be achieved with it.

The experiment

In some of the English language courses taught at the University of Padova, experiments are being carried out on the use of the speech analysis software Praat (freely downloadable from http://www.fon.hum.uva.nl/praat/) as a tool to aid pronunciation teaching/learning, with particular regard to prosodic features (see also Busà, 2008; and forthcoming). The aim of the experiments is to investigate the short-term and long-term benefits of the use of audio-visual feedback in pronunciation learning; the final purpose of the study is to implement a system that students can use in class or to study autonomously. The study is also aimed at investigating how non-native Italian intonation in English affects intelligibility.

For the experiment reported in this paper, 2 native (NS) English speakers (from Great Britain) and 8 native (NNS) Italian speakers (from the North-East of Italy) served as subjects. The subjects were asked to read aloud short English dialogues which were recorded and digitized using the speech analysis software Praat. Some phrases were extracted for comparison and analysis (see below). The study was designed to obtain preliminary data on the differences in intonation patterns of three sentence types (open questions, yes-no questions, and salutations) by English and Italian speakers. Through the comparison of the native and non-native sentences, the study aimed at getting evidence of how differences in intonation patterns may affect the intelligibility of the Italian-accented English speech.

Comparison of the intonation patterns in the NS and NNS productions and improvement after audio-visual feedback

Figures 1-9 compare the intonation patterns in one of the NS’ productions of an open question (‘What are you doing this evening?’, Fig. 1), a yes-no question (‘Are you going?’, Fig. 4), and a salutation (‘Bye!’, Fig. 7) with corresponding sentences produced by the NNs, before and after audio-visual feedback (Fig.s 2-3; 5-6; 8-9). The figures reproduce the type of visualizations of the speakers’ utterances pitch patterns and sound waves that were obtained with Praat. In each figure, the upper box shows the speech sound wave, and the lower box the corresponding pitch contour.
Fig. 1.: NS’ production of the sentence ‘What are you doing this evening?’

Fig. 2.: NNS’ production of the sentence ‘What are you doing this evening?’ before audio-visual feedback

Fig. 3.: NNS’ production of the sentence ‘What are you doing this evening?’ after audio-visual feedback
Figures 1-3 show a comparison of the production of the open question ‘What are you doing this evening?’ by a NS (Fig. 1) and a NNS before (Fig. 2) and after (Fig. 3) audio-visual feedback. It can be observed that the NS has a clearly falling intonation, with a prominence peak around the word ‘you’. By contrast, in Fig. 2, the NNS shows a pitch contour with three prominence peaks and a final rising intonation. After the feedback, the NNS’ production resembles the NS’, with a prominence peak on the word ‘you’ and a falling intonation.

Figures 4-6 show the production of the yes-no question ‘Are you going?’ by a NS (Fig. 4) and a NNS before (Fig. 5) and after (Fig. 6) audio-visual feedback. Here, again there is a clear difference in the intonation patterns produced by the NS and the NNS before the feedback: while the NS (Fig. 4) shows a marked rising-falling contour, with a pitch peak on the word ‘gOing’, the NNS’ utterance (Fig. 5) shows an intonation pattern which is characterized by a relatively level contour, a less prominent pitch peak than the NS, and placed in a different position in the utterances (at the onset of the vowel in ‘gOing’). On the other hand, the NNS’ production shows a noticeable improvement after audio-visual feedback (Fig. 6), with the production of a native-like rising-falling intonation contour, and a pitch peak at the end of the word ‘you’, following the prominence assignment rule of the NS.
Finally, Figures 7-9 show the visualizations of the utterance ‘Bye!’ The NS’ intonation contour (Fig. 7) is first rising and then level on a vowel segment that is 704 ms long. In comparison, the Italian’s utterance represented in Fig. 8 has an intonation pattern that is much more ‘flat’, i.e., with no clear contour or pitch peak. In addition, the Italian speaker’s diphthong in ‘bye’ is much shorter than the NS’, with a duration of 250 ms. After the audio-visual feedback (Fig. 9), the NNS’ utterance shows a clear improvement in both pitch contour and vowel duration, and it resembles much more closely that of the NS.

All the examples above prove that using audio-visual feedback helps learners to improve their L2 productions and get closer to the target utterance. Working with these visualizations provides learners and teachers with an immediate and easy-to-read image of the differences existing between the L1 and the L2, and does not necessarily require much knowledge of the phonological systems of the L1 and L2. In addition, most students enjoy the hands-on experience of working with their own language and discovering facts about it.

More investigations are needed to study the effects of this method in the long term.
6. Conclusions

The status of English as an international language calls attention to interlanguage communication and to the dynamics which may affect interactions between speakers from widely diverging linguistic backgrounds. An unclear pronunciation may be the cause of unsuccessful communication – by being detrimental to intelligibility, or be a source of discrimination and prejudice, and affect the L2-speaker’s social relations.

Research has shown that prosody plays a major role in affecting L2 speech intelligibility. It is used to signal information status in discourse, and to provide an interpretation to pragmatic meaning. Language teachers are called upon to provide learners with practice on English prosody, to help them communicate successfully. But if prosodic features are very important in discourse, they are also very hard to teach. Perhaps this is the reason why prosody is still largely underemphasized in English language programs in spite of its recognized role in the perception and production of L2.

A promising way to address the teaching of intonation has been through the use of speech visualizing technology, which has recently become widely available. Indeed, software like Praat allows users to record, listen and visualize their own speech, and compare it with native speakers’ productions. This allows learners to grow an awareness of their ‘distance’ from the target language and stimulates self monitoring and autocorrection.

This study has shown the kind of improvements that can be obtained with such speech-visualizing technology. The comparison of prosodic patterns in English as L1 and L2 has shown that the Italian speakers’ utterances differ from the native speakers’ in a number of significant features, and that audio-visual feedback helps the NNSs to improve their English prosodic patterns considerably, to the point that the latter come to resemble closely those of the NSs.

At this stage, more research is needed to investigate the long-term effects of the improvements obtained with the use of speech-visualizing technology. However, this study, in line with similar studies which are being carried out on a variety of languages, opens exciting perspectives for pronunciation teaching and learning. Prosody is becoming more accessible to the non expert, with concrete benefits for the learners. With an increased understanding of how prosody works for the L1 and the L2 speaker, teachers can help students to sound natural and be successful in their communication in English.
REFERENCES


Derwing, T., Rossiter, M., Munro, M.J. 2002. “Teaching native speak-
ers to listen to foreign-accented speech", *Journal of Multilingual and Multicultural Development*, 23(4), 245-259.


