REMARKS ON PROSODY AND FOREIGN ACCENT.
Giovanna Marotta (Pisa)

1. INTRODUCTION *

The studies on the acquisition of a second language have been increased in the last decades. However, quite often the research has been limited to the analysis of the errors made by the learners. Moreover, within the phonological domain, only the segmental aspects have been considered for a long time.

Two basic constraints emerge from the picture we can draw on this area of research:
- perception has been generally ignored;
- the role of prosody in L2 learning as well as in perceiving a foreign accent has not been taken into account.

Instead, we believe that prosody plays an important role in the process of L2 acquisition; in parallel, we think possible to disentangle the influence of segmental vs. suprasegmental features (i.e. the phoneme string vs. prosody) in what is perceived as a foreign accent.

Preliminarily, we might ask whether it is by chance that the term accent has been chosen to designate a foreign way of speaking. Some authors suggest a relevant plasticity of melody (e.g. Ladd 1996), which can easily be observed in infants in language acquisition. At the same time, it has been claimed that the typical categories of the mother tongue prosody are extracted very early, prior to the lexicon acquisition (cf. Ramus 1999, Ramus & Mehler 1999; Ramus, Nespor & Mehler 1999). Since prosody is essential and precocious in first language acquisition, we argue that in adults “perception and production routines are no longer very flexible and disrupt the processing of foreign languages in certain ways” (Dupoux et al. 1997: 407).

On the other hand, if we consider the linguistic atlases, it is noteworthy to observe that little or no space is usually devoted to prosody, which is a kind of an unexplored field even in dialectology. However, recent studies clearly show variation in prosodic features within the same language. For instance, the IViE project (cf. Grabe & Post 2002) has revealed how certain intonational patterns are specific to some English varieties. Prosodic variation diatopically constrained has been proved also for German (cf. Atterer & Ladd 2004) and Spanish (Sosa 1991, Prieto et al. 1995). For Italian too, some experimental research is now available showing different prosodic patterns with reference to different regions of the country (cf. Sorianello 2006: 118 ff.).

* It is a real pleasure to dedicate this paper to Lavinia, one of my best colleagues within the Faculty of Foreign Languages of Pisa University. The time spent with her in discussing how to organize our PhD program and solve relative problems has always been pleasant to me.
Furthermore, only a few studies analyze the role of prosody in the acquisition of a foreign language. The recent research attempts to elucidate the role of intonation and rhythm in characterising the global prosodic pattern of a linguistic variety (cf. Munro 1995, Dupoux et al. 1997, Magen 1998, Marcus & Bond 1999, Piske et al. 2001). However, two studies have to be pointed out: the first is the dissertation by Jilka (2000), which is completely devoted to the contribution of intonation to the impression of German accent in English speakers and English accent in German ones; the second is the very recent volume by Trouvain & Gut (2007), entirely devoted to non-native prosody.

To sum up, the relevant questions we may pose are the following:

a. since the prosody and grammar of a specific variety of language has been fixed in the mind of a child, the learning of new prosodic features, even different from the ones of his mother-language, will be more or less difficult for an adult with respect to the segmental features?

b. a speaker-listener exposed to a foreign accent will show the same or a different sensitivity towards the prosodic and segmental characteristics of L2?

If we consider the general picture of phonological studies concerning the acquisition of Italian as a second language, some basic elements emerge. First, the existing studies are recent and still not numerous; they concern the production of non-native speakers, with special attention to errors in the segmental aspects of speech, like the pronunciation of obstructive consonants (i.e. stops and affricates, both simple or geminates). Second, the different models of acquisition of L2 show the same scarce attention to prosody; consider for instance, the *Speech Learning Model* by Flege (1995, 1997, 2003), the *Perceptual Assimilation Model* by Best (1995), the *Ontogeny and Philogeny Model* by Major (1990, 1998, 2001). Finally, within the domain of these studies on Italian as L2, no specific room has been devoted to the perceptual side of the topic, probably because perception is judged as too difficult to be empirically investigated.2

However, we would like to underline the relevance of prosody in the perception of foreign accent, because foreign accent reflects segmental difficulties, but also suprasegmental ones. At the end of this paper, we will present the first analysis devoted to the perception of foreign accent by a group of Italian subjects listening to speech produced by speakers of different L1.

---

1 The first article on the topic is due to Bernini (1988); then, we have to wait for more than ten years later to find a series of papers devoted to the acquisition of Italian as a second language; see Giannini (1997), Giannini & Costamagna (1998), De Rosa & Schmid (2000), Costamagna (2003), Celata (2004), Mori & Barkat-Defradas (2005), Mori (2006), Sorianello (in press).

2 Up to now, a few papers analyze the prosody of speakers learning Italian as L2; see Gamal (2006; 2007); Missaglia (2007), Sardelli (2006; in press).
2. **Perception and Prosody**

The research about the perception of prosody has been traditionally focused on lexical accent of natural languages; intonation as well as rhythm have been ignored. This attitude is still present in the recent article by McQueen & Cutler (1997): metrical features are assigned a secondary role in the perception process guiding the lexical access, whereas segmental indexes are recognized as more relevant; for instance, vowels appear to be much more important than stress position in order to recognize a word.

Only recently intonation has been studied in depth from the side of perception. As a matter of fact, empirical research has shown that simple description of melodic patterns, even when based on robust acoustic data, it is not enough for the identification of elements which are responsible for the perception of prosody.

Intonation is a basic, essential element of the communicative process. Its intrinsic value is given by the whole amount of information, both linguistic and para-linguistic, which can be transferred from speaker to listener *via* melodic features. Quite surprisingly, it has been proved that even minimal melodic variations can be detected by listeners and perceived as distinctive features. Given the relative weight of each prosodic feature, recent studies on the perception of tonal structure of natural languages have shown that only pitch variations belonging to a definite range can be detected and decoded. For instance, C. Gussenhoven (2003: 759) said: “listeners are more inclined to hear unaccented words before the first pitch accent as accented as this pre-accentual stretch of speech is higher.” In other words, higher values of F0 in the first part of the sentence give a stronger and more robust perception of the prominent vowel than low or medium F0 values.

At the same time, it is necessary to consider the interaction between different parameters, both physical and perceptive. A special attention has to be devoted to the relationship between duration and F0 modulation in order to understand the relative role played by these cues for the perception of prominence: listeners perceive modulated vowels as longer; at the same time, a longer vowel is normally perceived as prominent.

From a theoretical point of view, the relationship between perception and production cannot be ignored. Recent research on mirror-neurons is clearly supporting the Motor Theory with new evidence: the connections between “to speak” and “to make” are becoming closer and closer, both

---

3 On the topic, we refer the reader to Gussenhoven (2003; 2004). As for Italian, segmental duration and scaling have been proved as relevant cues for discriminating different regional accents within Tuscany (Marotta & Sardelli 2003, Marotta, Calamai & Sardelli 2004); for instance, with respect to Florence in the areas of Pisa and Livorno, there is a stronger increase of stressed vowel duration with a wider frequency range; at the same time, on the perception side, longer vowels and higher modulation are clearly identified as distinctive features for Italian spoken in Pisa or Livorno (Calamai & Ricci 2005).
from the philogenetic and ontogenetic sides. However, according to Fowler & Galantucci (2005), the very true and relevant implications of this scientific discovery are not completely developed within the domain of phonetics and linguistics.

Another theoretical question which has a strong relevance concerns the status of the prosodic component in the grammar. For instance, we might ask whether the intonation should be considered as an analogic/continuous phenomenon or as a discrete/categorical one. The answer to this question has important consequences on the perception side too.

If prosody is not a hardware component of grammar, in non-tonal languages the perception of melody will be preferentially or exclusively holistic; therefore, a theory of intonation based on tonal movements will be preferred. On the other hand, if intonation belongs to grammar in all natural languages, the perception of prosody will be a distinctive and categorical process; as a consequence, a level theory will be chosen instead of a pattern based on pitch movements.

3. PROSODY IN L2 AND FOREIGN ACCENT

As already discussed in the introduction of the present paper, despite the strong increase of studies devoted to prosody, the contribution of prosodic features to the perception of foreign accent is still under-evaluated.

Is is not easy to exactly define the concept of ‘foreign accent’. Following Jilka (2000: 9), by “foreign accent” we refer to “a deviation from the generally accepted norm of pronunciation of a language that is reminiscent of another language, i.e. the speaker’s native language”. In general, only those deviations that are perceived as such by native speakers can be considered as instances of foreign accent. The deviations can be due to the interference between L2 (i.e. the foreign language) and L1 as well as to general cognitive principles related with Universal Grammar. Not surprisingly, research on foreign accent has so far concentrated on segmental aspects, whereas the role of prosody is a recent topic.

To capture the role played by prosodic features in the perception of foreign accent is an ambitious target, not only because of the great variation found in melodic and rhythmical patterns, but also since prosodic parameters are strictly interwoven; as clearly stated by Jilka (2007: 90), “several intonational features together would conspire towards a specific overall intonation characteristic”. However, despite the “conspiracy” of prosodic features, we believe worthwhile to look for the role played by prosody in this linguistic game.

Within the literature devoted to the role of prosody in both topics of L2 learning and perception of foreign accent, English appears to be the most studied language. Experimental data collected with reference to the acquisition of English show some typical prosodic patterns which are
independent from the specific L1 of speakers and which clearly contribute to the labelling of their speech as foreign; among them, the most relevant features seem to be a reduced frequency range (flat intonation), a different distribution of Pitch Accents, a less steep slope in the declination line (cf. Mennen 2007: 55).

We might wonder whether is possible to extend these features to the learning of other L2 languages than English. In our opinion, the prosodic categories of L1 must be taken into account in order to explain any possible transfer from L1 to L2. Following the common path shared by learners of L2, a beginner speaker will focus primarily on the lexicon of the language, then on morphosyntax, and only as a thirth step on phonetics and prosody (see for instance Giacalone Ramat 1988; 2003). This does not mean that phonological aspects of linguistic behaviour are less relevant than others, but phonetic accuracy seems to appear later on the competence of L2 with respect to lexicon and grammar. On the other hand, without any doubt phonetics and prosody are the levels of language more involved in the syndrome of foreign accent. As Izbedski (2008: 203) wrote, “the act of speaking a foreign language at all, or speaking it fluently, is a formidable task (...). Among the parameters of language, prosodic features appear crucial to achieve accent-free production, with correct prosody being seemingly more crucial than correct articulation or lexical confidence”.

In the specific case of intonation, to identify the features of foreign accent and the virtual transfer of prosodic categories from L1 to L2 is very difficult, because of the wide range of variation which is normally found in the inspection of the melodic contour in any language. In other terms, it is not clear which intonational cues are actual evidence of foreign accent and not simply deviations within the range of allowed variation. However, we believe that the basic criterion could be the appropriateness of the phonetic production in a given context (cf. Jilka 2000: 82).

4. A PERCEPTIVE OPPURE PERCEPTUAL ?? STUDY

The role of prosodic features has been recently approached in a study devoted to what is perceived when we cross the segmental and suprasegmental features in different language (cf. Boula de Mareuil, Marotta & Adda-Decker 2004).

For this purpose, we built a dozen sentences, 15 syllables on average, which are (almost) spoken in the same way in Spanish and Italian; e.g. ha visto la casa del presidente americano (“you/(s)he saw the American president’s house”). Two Italian speakers and two Spanish speakers read the obtained sentences three times in a soundproof booth with a high-quality microphone, using a DAT (input sampling rate of 48 kHz). The data were then transferred onto the computer with a sampling

4 For a general picture of the studies devoted to prosody in L2, see Marotta (in press).
frequency of 22.05 kHz and a 16 bit resolution, for further processing. Only one repetition per sentence was retained for each speaker. The speakers, who were all under 40 years old and graduated, had nothing to do with the experiment. They were instructed to articulate properly, without however making too many pauses other than the ones punctuation marked.

The perceptual stimuli were a set of comparable sentences of Italian and Spanish, which were modified both segmentally and prosodically, by making use of diphone speech synthesis.\(^5\) The pitch and duration parameters are then handled thanks to the PSOLA algorithm.\(^6\) With diphone synthesis, we could even play on the allophone inventory of Spanish and Italian, as well as other parameters too. In this study, we used both prosody transplantation and cross-language phonemic transcoding. After recording of native speakers of Spanish and Italian, their prosodic parameters were extracted, checked manually and applied on the diphone bases. Then, we obtained four different classes of stimuli:

1. Italian voice with an Italian prosody
2. Italian voice with a Spanish prosody
3. Spanish voice with a Spanish prosody
4. Spanish voice with an Italian prosody.

The perceptual experiment took place in a quiet room, through headphones. The listeners, all with normal hearing, were not urged to answer, but they could listen to each stimulus only once. They were instructed that they would listen to acoustically modified speech, coming from native speakers of Spanish and Italian. They were informed that the test sentences which were read in either language by the speakers could be spoken almost in the same way in Spanish and Italian.

Forty subjects (twenty Italian and twenty Spanish, all students) participated in the experiment; the perceptual test was conducted in Pisa and Barcelona. Since the identification test was based on a multiple choice, each listener had to judge and identify every stimulus in one of the four following categories.

- Italian voice with an Italian prosody
- Italian voice with a Spanish prosody
- Spanish voice with a Spanish prosody
- Spanish voice with an Italian prosody.

---

5. This method of synthesis relies on the concatenation of pre-stored phonetic units stemming from natural voices.
6. Energy is only normalised. As for pitch, it is defined for each phoneme by an initial target, a final target and possibly an intermediate target: one or two linear pitch movements are thereby associated to each phoneme. The pitch value of supposed unvoiced segments is equated to zero, and the initial pitch of each phoneme is connected to the final pitch of the preceding one, if any. See Boula de Mareuil, Marotta & Adda-Decker (2004) for further details.
The collected results show that for both groups of listeners (Italian and Spanish), the sentences agreeing in voice and prosody (e.g. Italian voice with Italian prosody, Spanish voice with Spanish prosody) are correctly recognised in the great majority of cases. As it was easily supposed, a sympathetic cooccurrence of prosody and segments guides the perception of the listeners.

However, in crossed sentences (Italian voice with Spanish prosody, Spanish voice with Italian prosody), prosodic features appear to be more relevant than segmental ones, since stimuli with the original segmental structure of Lx, but modified prosody of Ly are recognised more often as belonging to Ly than to Lx. The results are highly significant \((p < 0.01)\), and suggest a very important role of prosody in the perception of foreign accent.

In our experiment, prosody happened to play a stronger role as compared to segments, all other things being equal, in the task of recognition of language. However, this finding should be taken with caution, since it may be partially due to the use of diphone synthesis, which produces a less natural speech. If it is the case, it is interesting to replicate the experiment with a less degraded signal.

We may wonder if this experiment enables us to assess the contribution of segmental and suprasegmental factors to the perception of foreign accent, or to language identification. Indeed, we very partially treated the phonological transfer phenomena, which may occur during L2 acquisition. Following the \(SLM\) model, according to which production and perception remain subject to adaptation across the life span, we can even imagine that, in a longitudinal study, the place of prosody increases as long as the accent fades. Indeed, segmental errors may mask prosodic errors, since serious phonological errors can draw attention away from mere phonetic “errors”.

There is no question about the relevance of segmental factors in the identification of a speaker as foreign; rather, the point we would like to mark out is the parallel relevance of the prosodic cues, in both aspects of rithmical patterns and melodic profiles. The result emerging from our perceptual test seems to be particularly significant, since the two compared languages show similar phonetic and prosodic structures. We could expect, according to \(SLM\) model (cf. Flege 1995; 1997; 2003), that Italian and Spanish, being both syllable-timed, could be hardly discriminated and with a low degree of success; instead, prosodic features have emerged as sufficient elements for a right discrimination between the two languages. Therefore, prosody should arguably be implemented in \(SLM\) model as well as in the other theoretical models of L2 acquisition, since speakers of different L1 are sensitive not only to different segmental contrasts but also to different prosodic properties, such as accent or melodic contours.
5. **FURTHER RESEARCH**

A special issue concerns the weight of phonetic features on the perception of foreign accent in speakers having a very good competence of L2.

We might ask the following questions:

- do some features of foreigness still remain in the performance of good learners of a second language?
- if this is the case, are these features the same regardless L1 of the speaker or do they depend on the specific properties of the language?

In order to try to answer to these questions, we planned a series of perceptual tests, based on the speech as produced by foreigners having different L1 but a very good linguistic competence in Italian, due to the long stay in our country and the daily use of the language since many years. We expect to find out the relative relevance of segmental and prosodic features in the recognition of foreign accent and in L1 recognition.

In the pilot perceptual test we performed up to now, short tokens of Italian speech produced by speakers of different L1 (French, Spanish, German, English) having an optimal competence of Italian were judged by native Italians. The tokens were extracted from some conversations as well as from the reading of a newspaper article; therefore, our stimuli could belong both to spontaneous and read speech.

The subjects had to listen to the natural stimuli, one-by-one, and judge whether the speaker was a native Italian or a foreigner; if foreigner, they had to identify speaker’s mother-language among the four languages listed above. Eighty subjects (almost all students) participated in the experiment; more than half of them were of Toscan origin, the others of other regions of Italy. Each Italian listener has preliminarly to self-evaluate his familiarity with the four languages on a 3-point scale, going from 0 (no accent) to 2 (strong accent).

The results collected show that in the great majority of cases, listeners are capable of perceiving non-nativeness even in speech produced by foreigners having an optimal competence of Italian. However, to identify the mother-language of the speakers has been a task much more difficult. The English stimuli are the only ones showing a rather satisfactory success ratio, whereas low values of right recognition are associated to the stimuli produced by Spanish subjects, quite often classified as Italians. Furthermore, the speech by German speakers are frequently confused with the English ones, and sometimes, with the French ones too.

---

7 For our perceptual test, we selected short stimuli, in order to avoid the additional support due to the iteration of ‘foreign’ features; as a matter of fact, the perception of foreign accent increases as the phonetic sequence becomes longer (cf. Jilka 2007: 92).
It is easy to see that the degree of success in recognizing L1 of the speakers by Italian listeners is inversely proportional to the proximity between the two languages (L1 and L2): Italian and Spanish are hardly discriminated, whereas a German speaker can be confused with an English much more than with a Spanish.

Moreover, we found no correspondence between listener’s self-evaluation of his familiarity with the foreign language and performance in the perceptual test. English accent only shows comparable values of self-evaluation and actual perception,\(^8\) while for the other foreign languages we observed a more or less marked discrasy between the supposed level of familiarity with a foreign accent and the real identification of that accent in performing the experimental test. In other terms, the perception of foreign accent can be independent from the identification of the language spoken by a ‘foreigner’.

Another point of interest coming out from the results, concerns the difference between the two different styles of speech (i.e. spontaneous and read speech) we used in the test: quite surprisingly, listeners have succeeded in the identification of the mother-language of the speakers in read speech more than in spontaneous speech; see Tables 1 and 2 below.

<table>
<thead>
<tr>
<th>Read</th>
<th>English</th>
<th>French</th>
<th>German</th>
<th>Italian</th>
<th>Spanish</th>
<th>Simply foreigner</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>60%</td>
<td>10%</td>
<td>10%</td>
<td>8%</td>
<td>1%</td>
<td>12%</td>
</tr>
<tr>
<td>French</td>
<td>2%</td>
<td>51%</td>
<td>12%</td>
<td>19%</td>
<td>4%</td>
<td>13%</td>
</tr>
<tr>
<td>German</td>
<td>14%</td>
<td>16%</td>
<td>62%</td>
<td>0%</td>
<td>1%</td>
<td>7%</td>
</tr>
<tr>
<td>Italian</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>96%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Spanish</td>
<td>4%</td>
<td>12%</td>
<td>6%</td>
<td>24%</td>
<td>34%</td>
<td>21%</td>
</tr>
<tr>
<td>Mean value</td>
<td>16%</td>
<td>18%</td>
<td>18%</td>
<td>30%</td>
<td>8%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Table 1: Percentage of right answers in read speech

<table>
<thead>
<tr>
<th>Spont</th>
<th>English</th>
<th>French</th>
<th>German</th>
<th>Italian</th>
<th>Spanish</th>
<th>Simply foreigner</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>43%</td>
<td>5%</td>
<td>26%</td>
<td>8%</td>
<td>3%</td>
<td>15%</td>
</tr>
<tr>
<td>French</td>
<td>2%</td>
<td>22%</td>
<td>2%</td>
<td>45%</td>
<td>17%</td>
<td>12%</td>
</tr>
<tr>
<td>German</td>
<td>16%</td>
<td>28%</td>
<td>36%</td>
<td>8%</td>
<td>2%</td>
<td>10%</td>
</tr>
<tr>
<td>Italian</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>95%</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>Spanish</td>
<td>2%</td>
<td>8%</td>
<td>3%</td>
<td>32%</td>
<td>49%</td>
<td>6%</td>
</tr>
<tr>
<td>Mean value</td>
<td>13%</td>
<td>13%</td>
<td>14%</td>
<td>37%</td>
<td>14%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Table 2: Percentage of right answers in read speech

---

8 This datum is not surprising, since English is the foreign language young Italian people are more exposed to trough mass-media at present time.
The fact that native Italians could more easily recognize the L1 of the speakers if were listening to formal (read) speech can hardly be interpreted in the light of the *Ontogeny Philogeny Model* (OPM) as proposed by Major (2001): the stilistic corollar says that L1 increases in informal (= spontaneous) style. Therefore, the prediction is that the native listener can better judge where the foreign accent comes from in spontaneous speech, where the degree of self-control is lower. In our results, formal speech appears to be that more easily identified; this could be due to the fact that all the speakers we had have followed a guided path of learning Italian, trough school or courses of Italian; therefore, their competence is grounded on formal language.

Our pilot study allows us to point out the following aspects:

- on the production side, a speaker of L2 can very hardly loose the phonetic-prosodic *imprinting* of his L1;
- on the perceptual side, the sensitivity of native speakers towards their mother tongue is so fine-grained that the perception of even one feature of a foreign language - segmental or suprasegmental – is sufficient to trasmettere the idea of “foreign accent”.

We now wonder whether there exist a hierarchy between segments and suprasegmentals in the process of perception. Further research, crucially based on perceptual tests, could possibly answer to these questions. At the moment, we would like to conclude this article with the explicit aknowledgement that prosody is a factor non subsidiary in the perception of” foreign accent”. The awareness of the difficult task does not allow us to stop along this line of research, whereas it could become a new and robust stimulus to move on.
References


