

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Teaching the Prosody of Emotive Communication in a Second Language

*Anna De Marco*

## Abstract

The expression and the perception of emotional states in a foreign language represent a difficult task for the learners. One of the reasons is the fact that, more than other aspects related to speech, the expression of emotional states in second language requires full control of the prosodic resources that contribute to their realization. The aim of this chapter is to give an overview of the main tenets of the interface between prosody and pragmatic competence in L2 and in particular the expression and perception of emotions. The chapter will also outline some of the outcomes of the research in the field, focusing on experimental studies that have been conducted with learners of Italian as L2. The second part of the chapter will be devoted on the instructional practice aimed at developing the awareness of pragmatic-prosodic aspects of emotive communication in speech. Teaching practices such as a training focused on the expression of emotions (anger, joy, sadness, disgust, fear, and surprise) and video dubbing projects have proven to be useful tools to improve the performance of learners both in production and in perception of prosodic patterns of emotional communication.

**Keywords:** emotion, Italian L2, teaching practices, prosody, pragmatics

## 1. Introduction

This chapter is the result of the experience working with students of Italian L2 and, in particular, of the observation of their need to improve the competence of language use, that is, communicative competence. This competence involves different communicative dimensions even the most “hidden” and “subtle” ones but no less essential for the purposes of effective communication, such as the paraverbal (e.g., prosody) and nonverbal (facial expressions, gestures, postures) aspects which give form and meaning to our words and specify our utterances in ways that escape our awareness. Very often, when we find ourselves talking with a nonnative speaker, especially when we misinterpret each other, these paraverbal and non-verbal aspects come to light and call into question our complex competence as speakers because they represent a substantial part of our intentions and attitudes.

Only recently, in Italy, scholars have devoted themselves to the research on prosodic competence in L2, and the works dealing with this competence from a pedagogical perspective are even smaller. Indeed, emotional competence is one of the most overlooked aspects of communicative competence in teaching a native and a nonnative language. In a linguistic scenario that is ever more diversified in our

eyes, the conspicuous presence of students with different cultural and linguistic backgrounds draws the attention of educational institutions toward the need to adopt language teaching approaches that should be able to cope with the multilingual and multicultural complexity of our schools.

The first part of the chapter will be devoted to give an overview of the main tenets of the interface between prosody and pragmatic competence in L2, and it will introduce the reader to some of the outcomes of the research in the field, focusing on experimental studies that have been conducted with learners of Italian as L2.

The second part of the chapter will focus on the instructional practice to show how it is possible to improve the emotional competence of L2 learners through some teaching activities in the language classroom.

## 2. Theories on emotive speech

The study of the vocal expression of emotions is based on the analysis of complex structures that regulate the communication process: the voice constitutes a fundamental unit of measure within the emotional phenomenon. As D'Urso and Trentin explain: "the relationship between voice and emotion is based on the assumption that the physiological reactions typical of an emotional state, modifying the breathing, the phonation and the articulation of the sounds, produce appreciable variations in the acoustic indices detectable in the production of discourse" ([1], p. 58). The human voice, among its multiple potentialities, is able to convey different meanings and nuances through its own modulation. These prosodic modifications, which represent suprasegmental entities, merge with the segmental characteristics typical of each language. When we express an emotion, the meaning of our communicative act is conveyed not only by the voice and individual characteristics but also by the choice of the lexicon, by the way we organize our speech, and, of course, by the way we articulate sounds. As Poggi and Magno Caldognetto [2] maintain, there are four segmental resources that regulate the linguistic nature of emotional speech:

*Lexical resources:* That is, words (verbs, adjectives, adverbs, nouns) and interjections whose semantic and pragmatic content indicates a clear reference to specific states of mind (I feel guilty; I am sad, angry, etc.). In particular, interjections (damn, gosh) convey the emotional message without explaining its cause, which must necessarily be inferred from the context.

*Syntactic resources:* Emotional speech also influences the structure and organization of the utterances and is expressed through the alteration of the canonical order of the sentence. An example is the left dislocation, whose use implies the choice to focus a certain element of the discourse.

*Morphological resources:* The structure of words can undergo further transformation in emotional production. For example, it is very common to use diminutives and endearment words in relation to positive emotions, such as joy, and offensive words in the case of negative emotions, such as anger.

*Phonological resources:* In emotional speech, the articulation of sounds is altered by the subjective experience in a given context. Such alterations can be attributed, for example, to speaking while smiling or to other relevant aspects of nonverbal communication.

## 3. A cross-cultural perspective on emotive speech

Much of the research on emotional speech has attempted to identify common trends in the interpretation of vocal expressions of emotions at an intercultural

level. This approach should provide an indirect proof of the existence of universal (or specifically cultural) elements in emotional communication. However, there is a series of studies—less numerous—focused on the analysis of acoustic parameters and their variation at the interlinguistic level.

The research conducted by Scherer and other scholars has undoubtedly the merit of having conducted a systematic analysis, also examining cultures distant from each other. In this regard, we recall a study conducted in 2001 [3] which involved subjects from nine different countries (European, North-American, and Asians), in order to test their ability to recognize emotions expressed in another language (German). The body of emotional stimuli—conveyed through a nonsense sentence and pronounced by professional actors—included five emotions: joy, sadness, fear, anger, and disgust (along with neutral productions). The results showed an index of accuracy in identifying emotions equal to 50%. However, the percentage of correct decoding has not proven to be homogeneous in all the samples considered, since the Indonesian group has reached the lowest percentage, while the sample of German listeners recorded the highest percentage (in-group advantage), followed by the American group. The authors have attempted to explain the phenomenon, hypothesizing that linguistic closeness played an important role in the decoding process (language distance hypothesis). This accounted for the much higher percentage obtained by the British and the Dutch than the speakers of Romance languages (Italian and French) and above all with respect to the Indonesians (not belonging to the Indo-European family).

In addition to linguistic factors, sharing the same social codes and similar communication styles also plays an important role in the recognition of emotions. As Elfenbein and Ambady ([4], p. 204) explain: “It is possible that recognition is higher when the emotions are both expressed and perceived by members of the same cultural group” (see in-group advantage hypothesis). In this regard, Elfenbein and Ambady have suggested that the contact and the interaction between cultures, as well as the sharing of some cultural traits, can favor the decoding of emotions (cultural proximity hypothesis). On the basis of this hypothesis, “members of cultures who share ideas of individualism or collectivism, power structure, and gender roles, need to be more successful at decoding each other’s emotional expressions than members of cultures that are less similar” ([5], p. 409). In fact, according to the authors, the variability of the recognition indices is due not only exclusively to ethnic, geographical, or linguistic factors but also to the absence of mutual involvement between the cultures considered. To support this finding, some research has shown that the index of accuracy in the recognition of emotions at the cross-linguistic level is higher when the contact (even by telephone) between the cultures examined is greater [4]. Similarly, members of linguistic minorities within a given country would be able to decode more accurately the vocal expressions produced by members belonging to the majority culture than the latter in the reverse situation.

#### **4. Emotions in L2: state of the art**

Research on vocal expression and the perception of emotions in a second language are not numerous. The interest in this aspect of communication in L2 is relatively recent; however, the acoustic-perceptive investigations have highlighted some problems related to the management of emotional speech in a second language, with an emphasis on linguistic, cultural, and social factors.

As we pointed out above and as suggested by Dewaele, the typological and cultural distance from the target language can influence the ability of learners to manage emotional speech in L2: “SLA research shows that learners from ‘distant’

cultures experience significantly greater difficulties in identifying emotion in the L2 and in judging the intensity of that emotion than do fellow learners from 'closer' cultures with similar levels of proficiency" ([6], p. 375). In this regard, the first studies in L2 [7, 8] have shown that the competence linked to the vocal communication of emotions depends to a large extent on the degree of familiarity and knowledge of the cultural context of the target language [9].

Difficulties also exist on a purely prosodic level. A study conducted by Holden and Hogan [10] explored the perception of paralinguistic intonational trends in Russian and English learners, with the aim of evaluating "the emotional and attitudinal 'confusion' that may arise in the use of foreign intonation in L2" ([10], p. 70). The authors started from the consideration that the differences existing between the two linguistic systems in relation to pitch range are often the cause of misunderstanding in the interactions between Russian learners of English and native English speakers, highlighting the tendency of the latter to interpret intonation variations judged as neutral by native Russian speakers, as signs of anger or irritation.

For this reason, the groups involved were asked to evaluate different types of utterances (interrogative and exclamative) based on their emotional value, in order to verify "if there would be a significant change in the judgment of selection of 10 emotions and attitudes with a change of intonation, while keeping other phonetic factors constant" ([10], p. 67). The utterances were manipulated to reproduce the English intonation in Russian sentences and vice versa. Participants were also asked to evaluate the differences between original and manipulated productions.

Results showed that in relation to positive emotions, native English speakers judged the pitch of their own language "higher" than that of native Russian speakers in polar questions. Furthermore, native English speakers have shown themselves to be more sensitive to the intonation of the Russians than to the latter, who on the contrary have shown a lower sensitivity to the variations in tone of the English language. However, in relation to pitch range, both groups "reacted more negatively to the greater pitch range of Russian intonation in exclamations and yes-no questions" ([10], p. 84), giving the original statements in Russian a greater negative value than to the English originals, defined instead as more "passive" in the WH questions. As Ladd points out ([11], p. 94): "English and Russian listeners interpreted differences in sentences in broadly the same way, regardless of their specific cultural norms about what count about 'neutral'"

On the perception of paralinguistic intonation patterns in a second language, the study by Chen [12] focused on learners of Dutch of English origin (with a low level of competence) and on Dutch learners of English (with an advanced level of competence), in order to verify whether the prosodic variations linked to biological codes convey universally shared meanings on an emotional level. The adopted methodology was taken up by Chen et al. [13]. Even in this case, the groups involved have demonstrated a lingual-specific sensitivity to the paralinguistic intonation patterns produced in a foreign language (in this case a second language). Comparing the two studies, the author found out that "L2 English listeners differ from L1 Dutch listeners differ from L1 English listeners" ([12], p. 163). According to the author, the influence of L1 could have played an important role in the interpretation of prosodic variations:

First, there is strong evidence from both L2 English and L2 Dutch listeners that L1 transfer plays an important role in interpreting paralinguistic intonational meaning in L2, as in interpreting linguistic intonational meaning in L2. Also, there is an indication that L2 listeners may activate their knowledge about intonational universals embodied in the biological codes (in particular, Gussenhoven's Effort Code), which accounts for L2 Dutch listeners' native-like behavior in the perception of "emphatic" as signaled by pitch register.

In relation to the level of competence, Chen suggested that from the first phases of acquisition, the learners were able to catch some differences on the paralinguistic level between L1 and L2, to then improve this knowledge in the more advanced stages of the interlanguage.

The study by de Abreu and Mathon [14] focused on the perception of spontaneous emotional speech, which investigated the role of prosody in encouraging or hindering the recognition of emotions by a group of Portuguese French L2 learners. The stimuli consisted of some portions of spontaneous speech drawn from a corpus of prank calls made by a radio host against public institutions and service encounters, in order to trigger a reaction of anger in the victims. The linguistic content of the statements has been obscured by introducing white noise, “in order to keep only the prosodic information” ([14], p. 2). The subjects, native French speakers and Portuguese learners of L2 French with an intermediate level of competence, were involved in two tasks: the first aimed at the recognition of the statements of anger (decision task), and the second asked to evaluate the productions in terms of intensity (evaluation task). The recognition rates recorded for both groups (50% for the Portuguese and 62% for the French) led the authors to claim that “prosodic information represents enough information to allow subjects recognizing anger” ([14], p. 4). However, some differences emerged in the interpretations of anger utterances. In some cases, Portuguese learners evaluated utterances that contained pauses, repetitions, and errors as emotional productions other than anger (“not anger”) unlike the French. In other words: “it seems that Portuguese do not consider a sentence said with anger when there are disfluencies, unlike French listeners” ([14], p. 4). In terms of production, Komar [15] proposed a contrastive analysis of emotional speech produced by Slovenian speakers in English and by native English speakers. The elicitation of emotions was entrusted to the reading of a dialog in English. The results highlighted the tendency of Slovenes to use a “flat,” less dynamic intonation than natives, mainly due to the diversity of the two intonation systems. As Komar explains ([15], p. 4):

There are two main reasons for Slovenes sounding flat in English. First, the Slovenes produce the falling tones in a much narrower pitch range than the English, and second, the step up in pitch from the end of the falling pre-tonic segment and the beginning of the falling tone is significantly smaller compared to the step-up in pitch made by the English speakers.

These two factors, together with the state of anxiety and discomfort felt when one is not competent in a language, could be responsible for some communicative failures in interactions with native English speakers, who, according to the author, would be inclined to judge the less dynamic intonation of the Slovenes as a sign of disinterest and scarce participation (or rudeness). On a perceptive level, some studies have shown that access to the verbal and vocal content of the emotional message is not as automatic in L2 as it is in the native language [16–18]. In this regard, Chua Shi and Schirmer [19] have explored the process of integrating linguistic content with prosodic indexes in native and nonnative speakers of English of different origins. The stimuli consisted of a series of terms with positive, negative, and neutral valence, pronounced in a happy, neutral, or sad tone. In two separate experiments, the participants expressed a judgment about the emotional value of the stimulus based exclusively on the lexical content; later, they focused on the tone of the voice, thus excluding the linguistic level. The results showed some similarities between the two groups, in particular with regard to response times: “More importantly, both native and non-native listeners responded faster and more accurately when verbal and vocal emotional expressions were congruent as compared to when they were incongruent” ([19], p. 1376). In the case of incongruent verbal and vocal stimuli, the participants’ reaction was similar, regardless of the level of competence possessed.

In light of this, the authors have suggested that “the integration of verbal and emotional expressions occurs as readily in one’s second language as it does in one’s native language” ([19], p. 1376). Graham et al. [20] verified the ability of a group of native and nonnative English speakers to recognize some emotional productions made by native English speakers. The results showed that learners with a higher level of competence were not able to recognize more accurately the proposed stimuli. According to the authors, the ability to decode cannot be acquired in the absence of an intensive exposure to the cultural context of the target language or without a didactic intervention aimed at developing these skills.

Bhatara et al. [21] explored the relationship between competence in L2 and the ability to recognize positive and negative emotions in the target language, specifically American English. The subjects involved were of French origin, with a variable level of competence (established on the basis of the participants’ self-assessment). The perceptive experiment involved listening to some utterances in English made by professional actors and the decoding of the emotions expressed through a multiple-choice test. Afterward, the participants were asked to evaluate (on a scale) the pleasantness, the power, the alertness, and the intensity of the emotional stimuli. In addition to whole utterances, the stimuli also included simple vocalizations (or affect bursts). The results showed that learners with a high level of competence were not facilitated in the recognition of emotions, especially positive ones (joy, pride, interest, and relief). According to the authors, the increase in the proficiency of the English language could have compromised the perception and recognition of positive emotional states rather than facilitating it, concluding that: “increasing understanding of the L2 may be accompanied by a slight decrease in ability to understand subtle differences between positive emotions among other speakers” ([21], p. 11).

The correlation between the level of competence of the L2 (perceived by the learner) and the ability to recognize emotions therefore appears to be in doubt. However, recent studies [22] have shown that a high level of competence in the target language corresponds to greater accuracy in decoding positive audiovisual stimuli; moreover, cultural distance seems to influence recognition rates (as already reported in cross-linguistic studies).

In the Italian context, the studies that have investigated this particular aspect of communication in the acquisition process of the Italian L2 are still rather small. Recent studies [23] analyzed the emotional speech produced by Chinese learners of Italian with a high level of competence (C1 level of CEFR). A sample of native Italian speakers was also involved in the experiment. Emotions were elicited through a card task, i.e., a verbal interaction activity between two participants involved in a card game. As the authors explain, this procedure was adopted in order to elicit “emotional linguistic reactions in the players and arouse five different emotions (anger, anxiety, disgust, fear and surprise)” ([23], p. 82). The emotional productions collected were analyzed using Praat, taking into consideration the following parameters: duration and number of syllables, full pauses and silent pauses, and maximum and minimum values of  $f_0$ . Based on these measurements, time, articulation rate, and pitch range (in semitones) were calculated. The analyses also considered the presence of affect bursts.

In relation to native Italian, the results confirmed the well-known distinction between high and low activation emotions. The expressions of anger, fear, and surprise were characterized by high  $f_0$  values and a wide tonal extension; on the contrary the expression of disgust presented lower values, being a low activation emotion. The emotional speech of the Chinese did not show, however, this variability: “as a matter of fact,  $F_0$  height and tonal range are quite steady in the whole corpus. The only exception is represented by anger and fear that are expressed with

slightly higher values. These data seem to suggest that Chinese learners vary their pitch account to distinguish different emotional states in the case of native Italian speakers” ([23], p. 83). Production in nonnative Italian was generally characterized by a slowing of speech rate (due to the attempt to articulate single words with greater precision). The use of a modulated intonation and variable in the expression of emotions seems therefore to be a common tendency in learners of a second language. This data was also reported in some pilot studies [24, 25] conducted on Indonesian and Polish students, whose productions highlighted similar intonation contours in all the emotions investigated (joy, anger, sadness, fear, surprise, disgust, and neutral speech). Furthermore, the perceptive investigations conducted in these studies have suggested the hypothesis that the emotional speech produced by the learners in Italian L2 may be not very effective in terms of communication, leading to misunderstanding with native speakers [25].

#### **4.1 The role of transfer in Italian L2**

The role of L1 transfer in emotional speech in Italian L2 was explored in a recent pilot study [26], which involved native Italian speakers (three males and two females) and Russian (two females) and Persian (one male and one female) Erasmus students with a B2 level of competence. All the students involved in the study were following an Italian (grammar and communication course), within the same class, at the University of Calabria, and they had lived in Italy for 9 months at the time of the experiment. The elicitation of emotions has been entrusted to the reading of a text, which has been inserted a standard phrase: “It is not possible” (adapted by [27, 28]), translated into the three native languages. To favor the identification of the subject, they were invited to express the required vocal emotion and equivalent facial expression simultaneously. The participants had the opportunity to repeat their performances several times, in order to reduce any inhibitions or insecurities due to the presence of the microphone.

The learners were invited to express their emotions in their language and later in Italian. The parameters investigated were divided into two macro-categories: temporal (computed in milliseconds of the total duration of the target sentence and of the speech rate); intonation, i.e., the pattern and the values of the fundamental frequency ( $f_0$ ), initial frequency (onset), final frequency (offset), the average frequency, and the melodic excursion of the utterance (calculated in semitones). From the results emerged that in relation to the temporal parameters (total duration and speech rate), with equal number of syllables, all the emotional statements produced by the learners in their native language were shorter than those articulated in Italian L1, with the sole exception of the neutral utterance. This pattern has suggested the hypothesis that Russians and Persians, when they express themselves in Italian, fail to adequately control the elocution parameter, especially when the behavior of emotion differs significantly from that of the native Italian, thus determining emotional productions “different” from those expected. Indeed, it was observed that the emotional utterances produced by the learners in Italian L2 did not differ from those produced in their L1. The effect of transfer is, therefore, confirmed. In relation to the intonation parameters, the  $F_0$  excursion of the emotional utterances produced by the native Italian tended to be more extended than that produced in L2. The analysis of the utterances produced in the learners’ L1 showed, both in Russian and in Persian, a more reduced tonal excursion than native Italian, resulting in a clearly more monotonous production. In fact, a less dynamic intonation characterized the productions in Italian of both groups (to a greater extent, the statements produced by the Russian participants showed a narrower frequency range than the Persians, which instead managed to better modulate the intonation, in some

cases approaching the target language). In this case, it was difficult to attribute the lack of dynamism of the learners' speech to a real prosodic transfer, as, as already mentioned, such behavior can be justified by the discomfort felt when speaking a foreign language and even more so in contexts that invest the emotional sphere.

From a recent study [29], a heterogeneous picture emerged in relation to the L1 of the learners involved (Russians, Tunisians, and Spaniards). The acoustic verification of the emotional productions made by the students in their L1 and in Italian showed a clear evidence of transfer, especially in the speech of the Tunisians, who tended to reproduce in Italian intonational contours very similar to those of their native language. In the productions of the Russians and the Spaniards, instead, the contours, partially overlapping and rather monotonous, lost part of their auditory distinctiveness. According to the authors, the productions in L2, in some ways confused and very often not congruent on a pragmatic level, represent an indication of the difficulty faced by learners when they realize in a different language a complex paralinguistic phenomenon such as emotions.

## 5. Teaching to express and perceive emotions in L2: a prosodic training

In this section, we will deal with the implementation of a training for the classroom practice. Learners involved in this study are Persian and Russian students who participated in the research described in the previous paragraph. The structure of the training is partly structured following the didactic approach of the task which is defined by Nunan as follows:

*A piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning and in which the intention is to convey meaning rather than to manipulate form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right with a beginning, middle and an end. ([30], p. 4)*

During the lesson, students concentrate on the meaning and not on the linguistic form and are involved in tasks that simulate or reproduce what happens outside the school context. However, reflection on the language may arise during the work in class and therefore from the need of the learners to reach the communicative objective they have set for themselves. Given the experimental nature and the prosodic focus of the activities, we have planned tasks of different nature that the students have prepared to face a reflection on the paraverbal and non-verbal elements of the emotions that were the object of the training. The first activity involving learners on the theme of emotions is a meta-pragmatic activity. The next phase of the experimentation focused on the presentation of linguistic and communicative inputs of an emotional nature. The learners viewed some film clips showing examples of emotional states (verbal and non-verbal) and were involved in role-play activities. Other non-verbal activities prepared the learners for the actual task which consisted in the planning and realization of a dubbing. Previous studies [31–34] have shown that video technologies are a didactic tool with a strong motivational component, able to involve and stimulate learning in an environment free from anxieties and apprehensions. To assess the effectiveness of the training, we adopted a cross-analysis of learners' production and perception skills—through acoustic analysis and auditory tests—before and after the training. This allowed us to highlight any improvement following the didactic intervention and to formulate further hypotheses. The auditory tests and the acoustic analysis were carried out according to the procedures described above.

Now let us see in detail the articulation of the training which took place over 4 weeks and was divided into four phases.

#### Phase 1: Meta-pragmatic activities: brainstorming

The aim of the initial phase was to provide language learners with space for a reflection on the language. The meta-pragmatic activity was aimed at an exchange of impressions and information on the modalities of expression of emotions in the learners' countries of origin in order to identify the main criticalities experienced by the learners in the interactions with the natives. This reflection also allowed learners to check the correspondence between the emotional labels proposed in Italian and those used in their L1. In this phase emerged the first criticality experienced by the learners of both groups in the interactions with the natives, due to the Italian habit of amplifying the volume of the voice, during the verbal exchange. Especially in Persian culture, this vocal variation is interpreted as a display of anger. This difference is often the origin of misunderstandings and misinterpretations.

#### Phase 2: Exposure to linguistic inputs and focus on prosodic aspects

This phase, which we can consider a phase of preparation for the actual tasks, has been focused on exposure to linguistic input, taken from Italian films and from the corpus of emotions produced by the natives. In particular, in the case of films, other non-verbal aspects (facial expression, gestures, postures, distance between interlocutors, etc.) involved in the expression of emotions were also identified. Subsequently, we gave the students a description of the mode of expression from a vocal point of view, focusing on the prosodic features [28, 34]. For example, as far as anger is concerned, we have given some information on the physiological alteration (irregular breathing, increased heart rate, etc.) and on the effects observable in speech (sustained speed, increase in intensity, etc.). The objective was to develop the awareness of the psychophysical changes and the effects on the vocal aspects triggered by the emotional states considered above. One of the activities proposed to the learners, who were working in a multilingual environment, was the vision of the clips without the audio, after which they were asked to guess the messages and then to try to realize the acoustic characteristics of the emotional expression conveyed by the messages themselves.

#### Phase 3: Non-verbal communication practice

The third phase of the training included the performance of some exercises related to non-verbal communication. Learners were divided into pairs and placed back to back. Afterward, they had to reproduce an emotional state by repeating the sequence of numbers from 1 to 4 (to avoid any interference of a semantic nature), letting the other person guess the reproduced emotional state. Some emotions have been expressed more easily (e.g., anger, sadness, and joy); others, however, have raised some difficulties (e.g., disgust and fear), both in production and recognition. In general, learners have learned to "play" with the intonation and the voice, managing to capture the emotional nuances, beyond the linguistic content of the message.

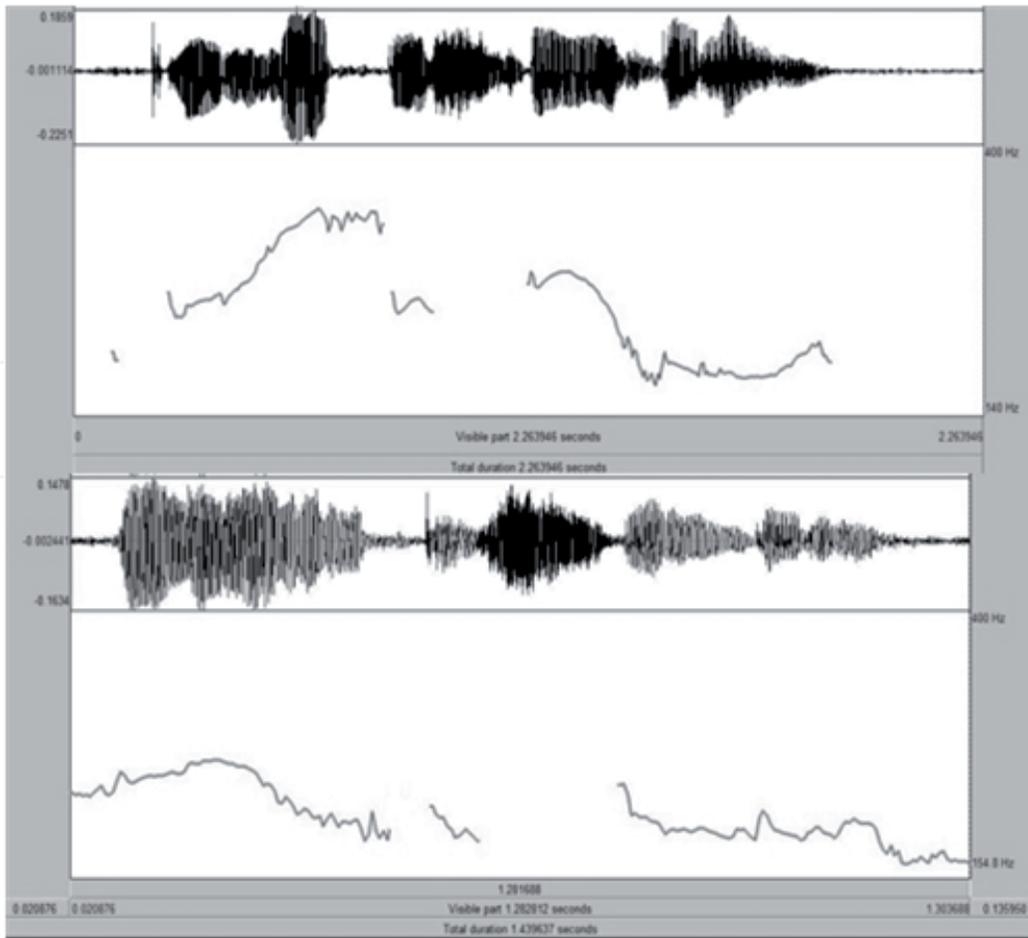
Learners were then involved in a series of playful activities such as interpreting certain dialogic parts characterized by a specific emotional context (e.g., an unexpected encounter between two friends, in which emotional states emerged, i.e., joy, sadness, anger, etc.) and then improvising a similar dialog resorting to some expressions used by the protagonists (see Appendix). The theatrical component of these activities gave the students the opportunity to use other non-verbal components during simulations. In order to enhance the focus on prosodic aspects, students carried out shadowing exercises, in which they were asked to listen to the emotional stimulus produced in Italian and to replicate it (in the shortest possible time) in terms of rhythm and intonation. In this last phase, learners had the opportunity to put into practice the suggestions provided in the previous phase.

#### Phase 4: The dubbing task

The actual task consisted in the realization of the dubbing. In our study, we selected some scenes (see Appendix) taken from an Italian film, *L'ultimo bacio* ("The Last Kiss" by Gabriele Muccino), containing specific emotional contexts (anger, sadness, joy, fear, and disgust). This phase consisted, in turn, of several moments in order to elaborate the task. During the preparation phase, we invited learners to carefully observe the scenes, initially projected without the audio, and to focus on the facial expressions and the gestures of the actors. Each learner proposed a personal interpretation of the video in an emotional key and tried to guess which of the primary emotions was expressed in the video. The same activity was replicated following the projection of the film with the original audio. This activity encouraged learners to exchange their impressions with respect to the hypotheses that they had previously formulated. The students were then divided into pairs, and each of them was assigned a scene to be interpreted, and they were asked to write the relating script. In this phase, the learners have been given a description of the emotions from a vocal point of view (speech rate, voice volume, intonation, etc.), analyzing together with the learners the characteristics of the actors' speech and its variability according to the emotional context. Later, the students were invited to individually and collectively experiment the lines of the dialogs assigned to each of them. Their interpretative effort focused on the emotions that were realized paying attention to gestures, facial expressions, and composite vocal modulations. Shootings were done with a digital camera, and audio recording was done using a professional recorder with internal microphone, to facilitate the dynamism of the scenes. At the end of the shoot, the material obtained was processed through a video editing software (Windows Movie Maker). With the active collaboration of the students, the assembly of the scenes was carried out; finally, the films obtained were screened in class, and the projection was followed by a meta-pragmatic discussion on the use of the voice and the body in the expression of emotions.

#### Acoustic analysis and auditory assessment: results

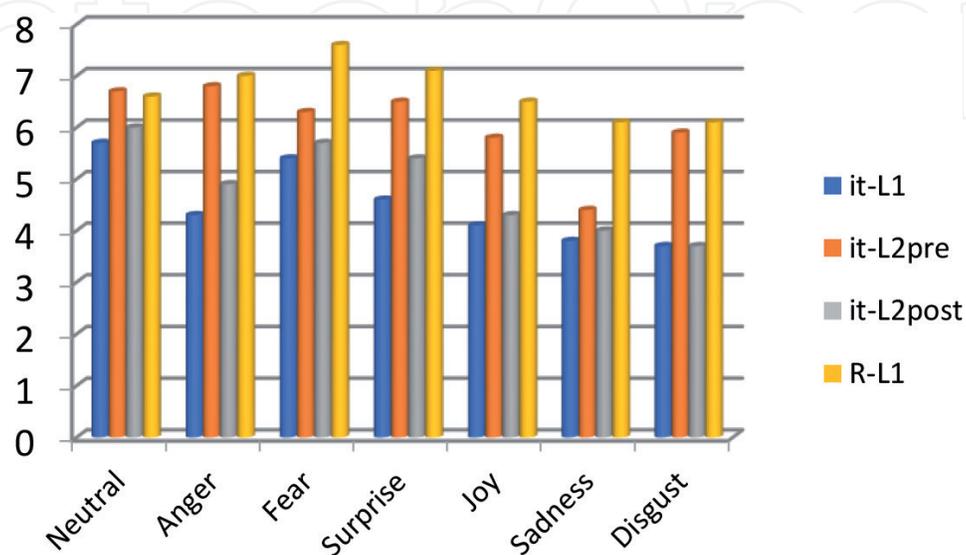
The data collected in the first and last phase of the training were subject to an acoustic analysis. Furthermore, before the training began, an auditory test was presented to both groups containing the six considered emotions (together with the neutral speech), performed by a native Italian speaker, in order to highlight any initial critical points. The hypotheses formulated at the outset were confirmed by the results obtained, highlighting the presence of concrete difficulties in the management of emotional speech by the subjects involved, both in terms of production and in terms of their auditory recognition. The acoustic analysis revealed some elements worthy of reflection. First of all, among all the investigated parameters, the most resistant index seems to be the intonation contour. Moreover, the analysis of learners' emotional speech in their L1 confirmed the hypothesis of the influence of the prosodic structures of the native language on those of the target language. In the melodic contour the emotions produced in L2, different phenomena of prosodic transfer can be noticed; the native language of the learners acts as a powerful filter that slows down, and sometimes blocks, the acquisition of the new intonation patterns. For example, the intonation contour of sadness realized by the Russian learners in L2, at least before the training, is similar to the one they produce in their native language, while the melodic excursion related to surprise never reaches hearing adequacy in the L2 of Persian learners, being less modulated than that of native Italian, but approximate to the melodic excursion of native Persian. In **Figure 1** we report by way of example the intonation contour relating to the sadness produced in Italian by a Russian learner. Before the training, the contour, placed in a wide tonal space, settles on high F0 values (EM: 11 ST); after the training we can see not only a re-dimensioning of the melodic excursion (7.7 ST) but also the correct positioning



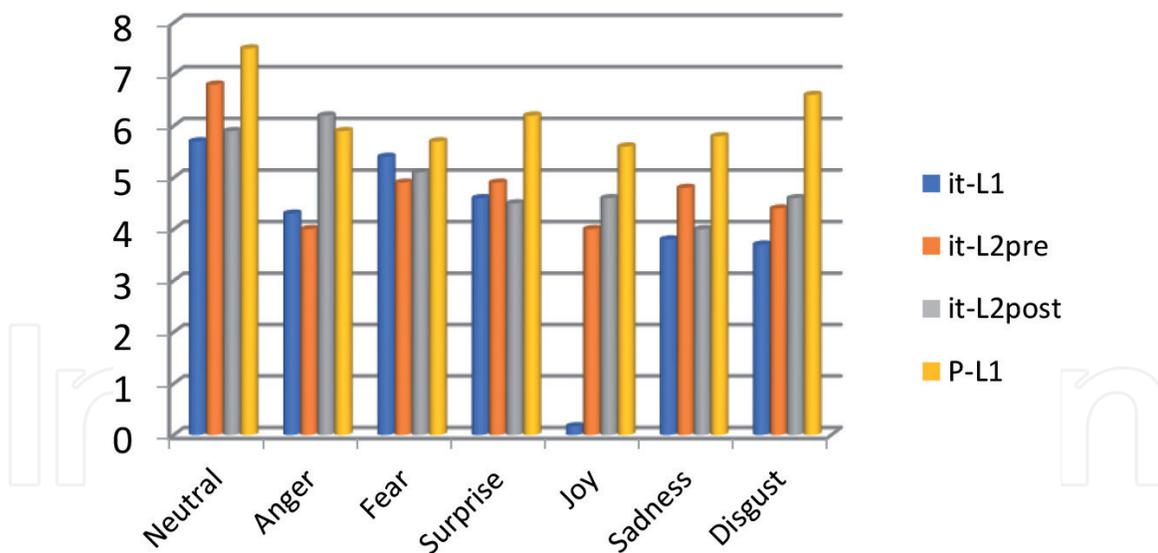
**Figure 1.**  
 Waveform and intonation contour of the sentence *non è possibile* (“it’s not possible”) conveying sadness produced by a Russian learner before the training (upper figure) and after the training (lower figure).

of the final contour; in this last case, the hearing result is obviously improved. However, it should be pointed out that emotional intonations only partially improve, and not all, under the action of training.

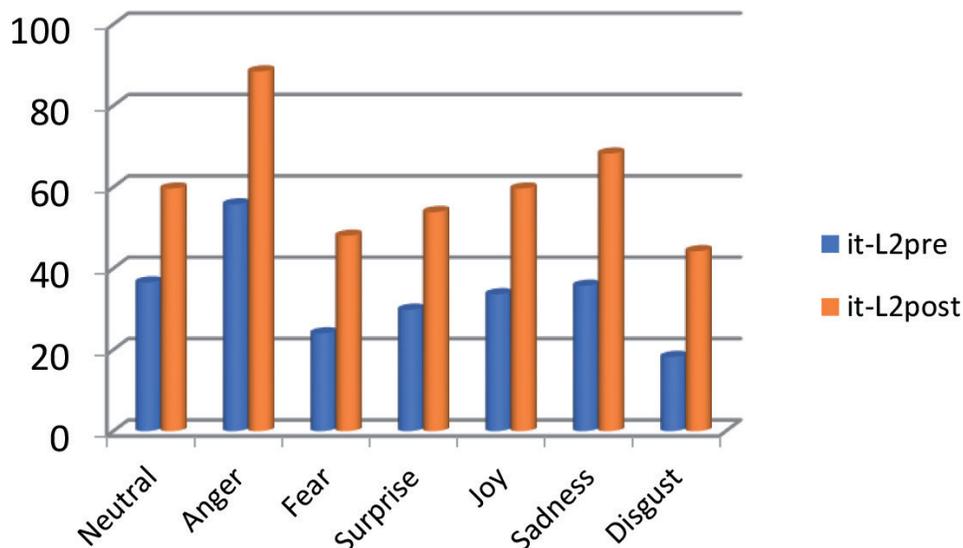
Learners are able to better control temporal parameters, duration, and speech rate and consequently to adapt them to the new linguistic model (Figures 2 and 3).



**Figure 2.**  
 Speech rate. Comparison of Russian learners’ productions before (*it-L2pre*) and after (*it-L2post*) the training. The figure also shows the variation of speech rate in learners’ L1 (*R-L1*).



**Figure 3.** Speech rate. Comparison of Persian learners' productions before (*it-L2pre*) and after (*it-L2post*) the training. The figure also shows the variation of speech rate in learners' L1 (*P-L1*).



**Figure 4.** Percentages of mean recognition accuracy of emotional sentences produced by learners before (*it-L2pre*) and after (*it-L2post*) the training.

These two parameters therefore show greater ductility, and presumably these are the ones that contribute to improving the auditory responses of the subsequent experimental phase. For example, the speech rate, anger, surprise, joy, and disgust found in Russian learners before teaching training show a strong conditioning by their native language, an effect that is, however, reduced after the training phase (**Figure 2**). The overall improvement was proven by the decoding of native speakers<sup>1</sup>: learners' emotional productions after the training have been recognized by native speakers much more than the pre-training productions (**Figure 4**). On the level of auditory

<sup>1</sup> The vocal productions of the learners (before and after the training) and those of an Italian speaker were included in an online auditory test and submitted to a sample of 26 Italian listeners (12 men and 14 women, aged 18–66 years, mean age: 35 years). The initial page of the test provided a brief introduction on the object of the research and a series of indications on the modalities of execution. The test consisted in listening to the stimulus and selecting the response within a set of options (joy, anger, disgust, fear, sadness, surprise, and neutral). The listeners had the opportunity to listen to the audio file several times and to add comments and personal opinions. Stimuli were played randomly.

recognition, the learners have shown a marked improvement in their decoding skills, correctly recognizing the emotions that initially caused major problems (although the Persian group encountered some difficulties).

## 6. Conclusions

The results of this study lead us to some more general considerations regarding possible interpretations on the acquisition of prosody in L2, including some observed transfer phenomena. Although the linguistic distance between L1 and L2 potentially constitutes an obstacle to the acquisition of new prosodic structures and their decoding, cultural distance plays in many cases a very important role in the elaboration of the perception and production of emotional states. From a verbal report conducted with the Russian and the Persian learners, extremely different communication profiles emerged with respect to the target language, and their perception of emotion is interpreted and adapted according to the cultural norms. In particular, the strong discomfort of the Persians arises from a different interpretation of the non-verbal signals of the Italian natives; for example, in Persian culture, prolonged eye contact is often decoded as an act of defiance; the voice volume of the conversation tends to remain rather calm, and the use of gestures is moderate. In the Italian culture, a communicative exchange is instead based on maintaining a visual contact, interpreted as a sign of attention, sincerity, and participation; moreover, in a normal conversation, participants continually resort to gestures and facial expressions to emphasize the message, and the volume of the voice adapts to the general emphasis and tends to rise considerably. All this is interpreted by Iranian speakers as an act of aggression; on the other hand, the Italians themselves could interpret the non-verbal behavior of the Persians as signs of coldness and low emotional participation.

These are therefore two essentially different communicative styles that are often the cause of misunderstandings due to the reciprocal tendency to interpret emotional speech according to its own cultural and linguistic parameters, considering them universally shared. As we have observed at length in the course of our discussion of the emotional phenomenon, vocal productions that manifest certain emotional values in a given culture can, in another, characterize contexts that are not specifically emotional or neutral. The difficulties encountered by the Persian learners in the interpretation of some verbal and non-verbal behaviors of Italians amply demonstrate this aspect. We therefore hypothesize that the transfer is not only of a prosodic nature, but that it has deeper roots, closely linked to the behavioral patterns that we have learned to manage and decode since our infancy and that are part of an adult's cultural and emotional memory. In this perspective, the management of emotional speech in an L2 is not limited to the correct use of the intonation structures of the target language, but involves deeper levels, perhaps less sensitive to the educational intervention. Moreover, the transfer acts as a sort of filter that the learner uses to elaborate the hypotheses on the target language being not the mere passage of a structure from one language to another. The development of an emotional competence in the target language represents, as we know, a delicate and personal path. However, we believe that focused instruction, similar to what we have proposed, can positively contribute to the process of interaction between very different linguistic and cultural systems.

As a matter of fact, the training carried out in this study, beyond the cultural and typological distance existing between the native languages involved in the experiment, has nevertheless produced positive effects. A long-term verification

still remains to be carried out, in order to determine whether the phenomena under investigation have been actually acquired.

## Thanks

I would like to thank Emanuela Paone for the careful and extensive revisions, suggestions, and the helpful comments on the chapter.

## Appendix

Screenplay: “The Last Kiss” by Gabriele Muccino, 2000.

La Confessione

The confession

Anna irrompe nello studio del marito psicologo (Emilio), mentre questi è impegnato con una paziente.

Anna breaks into the studio of her husband, the psychologist (Emilio), while he is engaged with a patient.

A.: *Eccoci qua!*

A.: Here we are!

E.: *Cos'è successo?*

What happened?

A.: *Secondo te niente di importante, vero?*

A.: In your opinion, nothing relevant, isn't it?

E.: *Sto lavorando.*

I'm working.

A.: *Lo vedo.*

I see it.

E.: *E allora esci e ripassa tra quaranta minuti, quando avrò finito.*

And then go out and come back in 45 minutes, when I'm done.

A.: *Ti devo parlare.*

I have to talk to you.

E.: *E non potresti parlararmi tra quaranta minuti o stasera, quando torno a casa, di grazia?*

So couldn't you talk to me in 45 minutes or tonight, when I get home?

A.: *NO!*

NO!

E.: (rivolgendosi alla paziente) *È mia moglie, non si preoccupi.*

(turning to the patient) She's my wife, don't worry.

A.: *Te la devo dire subito questa cosetta, di grazia! (arrabbiata)*

I have to tell you this *little thing* right away, goodness! (angry)

(Emilio si alza e si avvicina minaccioso alla moglie)

(Emilio gets up and approaches her wife threateningly)

A.: (sarcastica) *Dio, che paura! Sto provocando una reazione!*

(sarcastically) God, what a fear! I'm provoking a reaction!

E.: *Ma tu hai bevuto, eh?*

But you've been drinking, haven't you?

A.: Ah, now I've been drinking, huh?

Tonight, I leave you, my dear, you understand what

I said? I'll leave you forever!

E.: *Adesso fuori! Fuori! (urla)*

Out now! Out! (yell out)

A.: TI HO TRADITO! Hai capito che ho detto? TI HO TRADITO!  
Arrivederci e tante scuse per il disturbo  
(rivolgendosi alla paziente).  
I HAVE BETRAYED YOU! You understand what  
I said? I CHEATED ON YOU!  
Goodbye and many apologies for the trouble  
(addressing the patient).

IntechOpen

IntechOpen

### **Author details**

Anna De Marco  
University of Calabria, Italy

\*Address all correspondence to: [demarco.anna@gmail.com](mailto:demarco.anna@gmail.com)

### **IntechOpen**

---

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] D'urso V, Trentin R. *Introduzione Alla Psicologia Delle Emozioni*. Bari: Laterza; 1998
- [2] Poggi I, Magno Caldognetto E. Il parlato emotivo. Aspetti cognitivi, linguistici e fonetici. In: Leoni FA, Cutugno F, Pettorino M, Savy R, editors. *Atti del Convegno Il parlato italiano*. Napoli, Cd-Rom: D'auria Editore; 2004
- [3] Scherer KR, Banse R, Wallbott HG. Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*. 2001;32(1):76-92
- [4] Elfenbein HA, Ambady N. Is there an in-group advantage in emotion recognition? *Psychological Bulletin*. 2002;128(2):243-249
- [5] Thompson WF, Balkwill LL. Decoding speech prosody in five languages. *Semiotica*. 2006;2006(158):407-424
- [6] Dewaele JM. Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *The Modern Language Journal*. 2005;89(3):367-380
- [7] Rintell E. But how did you feel about that? The learner's perception of emotion in speech. *Applied Linguistics*. 1984;5:255-264
- [8] Rintell E. That's incredible: Stories of emotion told by second language and native speakers. In: Scarcella R, Andersen E, Krashen S, editors. *Developing Communicative Competence in a Second Language*. Heinle and Heinle Publishers: Boston; 2011. pp. 75-94
- [9] Kim S, Dorner LM. "I won't talk about this here in America": Sociocultural context of Korean English learners' emotion speech in English. *L2 Journal*. 2013;5(2):43-67
- [10] Holden K, Hogan J. The emotive impact of foreign intonation: An experiment in switching English and Russian intonation. *Language and Speech*. 1993;36:67-88
- [11] Ladd DR. *Simultaneous Structure in Phonology*. Oxford: Oxford University Press; 2014
- [12] Chen A. *Universal and Language-Specific Perception of Paralinguistic Intonational Meaning*. Utrecht: LOT; 2005
- [13] Chen A, Gussenhoven C, Rietveld T. Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech*. 2004;47(4):311-349
- [14] De Abreu S, Mathon C. Can you hear I'm angry? Perception of anger in a spontaneous French corpus by Portuguese learners of French as a foreign language. In: *Proceedings of the Phonetics Teaching and Learning Conference (PTLC)*; 27-30 July 2005; University College of London: London. p. 2005
- [15] Komar S. The impact of tones and pitch range on the expression of attitudes in Slovene speakers of English. In: *Proceedings of PTLC 2005*; 27-30 July 2005. London: University College London; 2005
- [16] Anooshian JL, Hertel PT. Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*. 1994;8:503-554
- [17] Harris CL. Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development*. 2004;25:223-247

- [18] Harris C, Gleason JB, Aycicegi A. When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In: Pavlenko A, editor. *Bilingual Minds: Emotional Experience, Expression, and Representation*. Clevedon, UK: Multilingual Matters; 2006
- [19] Chua Shi M, Schirmer A. Perceiving verbal and vocal emotions in a second language. *Cognition and Emotion*. 2011;**25**:1376-1392
- [20] Graham CR, Hamblin AW, Feldstein S. Recognition of emotion in English voices by speakers of Japanese, Spanish and English. *IRAL-International Review of Applied Linguistics in Language Teaching*. 2001;**39**:19-37
- [21] Bhatara A, Laukka P, Boll-Avetisyan N, Granjon L, Effenbein HA, Bänziger T. Second language ability and emotional prosody perception. *PLoS One*. 2016;**11**(6). DOI: 10.1371/journal.pone.0156855
- [22] Lorette P, Dewaele JM. Emotion recognition ability in English among L1 and L2 users of English. *International Journal of Language and Culture*. 2015;**2**(1):62-86
- [23] Maffia M, Pellegrino E, Pettorino M. Labeling expressive speech in L2 Italian: The role of prosody in auto- and external annotation. In: Campbell AW, Gibbon D, Hirst D, editors. *Proceedings of the 7th International Conference on Speech Prosody*; Berlin, 20-23 May 2014; Speech Prosody Special Interest Group (SPROSIG). Urbana: Illinois. 2014. pp. 81-84
- [24] De Marco A, Paone E. The acquisition of emotional competence in L2 learners of Italian through specific instructional training. In: Gesuato S, Bianchi F, Cheng W, editors. *Teaching, Learning and Investigating Pragmatics*. Cambridge: Cambridge Scholar Publishing; 2015. pp. 443-470
- [25] De Marco A, Paone E. Uno studio sui correlati acustici delle emozioni vocali in apprendenti di italiano L2. In: Elia A, Iacobini C, Voghera M, editors. *Livelli di analisi e fenomeni di interfaccia*. Atti del XLVII Congresso Internazionale SLI. Roma: Bulzoni; 2016
- [26] De Marco A, Sorianello P, Paone E. L'acquisizione delle emozioni nell'italiano non-nativo: Un percorso didattico longitudinale. In: Gudmundson A, Álvarez López L, Bardel C, editors. *Romance Languages Multilingualism and Language Acquisition*. Frankfurt: Peter Lang; 2017. pp. 57-85
- [27] Anolli L, Ciceri R. La voce delle emozioni: verso una semiosi della comunicazione vocale non-verbale delle emozioni. Milano: Franco Angeli; 1992
- [28] Anolli L, Wang L, Mantovani F, De Toni A. The voice of emotion in Chinese and Italian young adults. *Journal of Cross-Cultural Psychology*. 2008;**39**(5):565-598
- [29] Sorianello P, De Marco A. Sulla realizzazione prosodica delle emozioni in italiano nativo e non nativo. In: Savy R, Alfano I, editors. *La fonetica nell'apprendimento delle lingue*. Milano: Officinaventuno; 2016. pp. 155-178
- [30] Nunan D. *Task-Based Language Teaching*. Cambridge: Cambridge University Press; 2004
- [31] Brooke S. Video production in the foreign language classroom: Some practical ideas. *The Internet TESL Journal*. 2014;**10**(10). Available from: <http://iteslj.org/Techniques/Brooke-Video.html>
- [32] Dubreil S. When students become directors: Redefining the role of the learner in the foreign language classroom. In: Lomicka L, Cooke-Plagwitz J, editors. *Language Instruction: Teaching with Technology*. Vol. 1. Boston: Heinle; 2003. pp. 129-137

[33] Burston J. Video dubbing projects in the foreign language curriculum. *CALICO Journal*. 2005;23(1):79-92

[34] McNulty A, Lazarevic B. Best practices in using video technology. *Teaching English with Technology*. 2012;12(3):49-61

IntechOpen

IntechOpen