

Lexical Stress Training of German Compounds for Italian Speakers by means of Resynthesis and Emphasis

Maria Paola Bissiri ^{a/b} Hartmut R. Pfitzinger ^a Hans G. Tillmann ^a

^a Institut für Phonetik und Sprachliche Kommunikation, University of Munich, Germany

^b Dipartimento di Scienze dei Linguaggi, University of Sassari, Italy

{mariapa|hpt|tillmann}@phonetik.uni-muenchen.de

Abstract

We propose a new method for teaching Italian speakers to pronounce German compounds. This method gives as feedback the prosodically corrected version of an utterance in the learners' own voice adding emphasis to the stress position, in order to help the learners better recognize the correct pronunciation and identify their errors. In our study, twelve Italian speakers with German as L2 pronounce German compounds in isolation and/or embedded in longer utterances. A first group of speakers receives two kinds of feedback: i) their own utterances corrected in their prosodic parameters pitch-contour, local speech rate, and intensity, and ii) a second corrected version, whose prosodic parameters correspond to an utterance pronounced by a native speaker putting emphasis on the stressed syllable. The second group of speakers receives two kinds of feedback as well: i) the traditional feedback, consisting in the utterances of a native speaker, and ii) the utterances of the native speaker putting emphasis on the stressed syllable. Perception results show that the three kinds of feedback involving emphasis and/or synthesis are more effective than the traditional method, but that there is no significant difference between those three kinds of feedback. We expected the feedback in the learner's own voice with emphasis to be more effective than the same feedback without emphasis. A possible reason could be that several Italian speakers imitated the emphasized version, causing the evaluators to judge the word stress as correct but assigning it a lower score than if the word had not been overstressed.

1. Introduction

The pronunciation of German compounds is difficult for Italian speakers because they tend to stress the second part of the compounds instead of the first. An example are separable prefix verbs: the prefix has always to be stressed, but Italian speakers tend to stress the verb instead. Even advanced Italian learners of German living in Germany for many years may make this kind of mistake systematically. A reason could be that Italian speakers have difficulties in perceiving the actual stress position in words pronounced by German native speakers. Their perception of stress might rely on different acoustic cues from the Germans.

In the Italian language, duration has been demonstrated to be the strongest acoustic cue (Bertinetto 1980; Bertinetto 1981). Since Italian speakers rely on this parameter, this could lead them to perceive the second compound part as being stressed instead of the first. In a study conducted by the first author, wrongly stressed German compounds pronounced by Italian speakers were corrected by copying F0 contours and durations of the segments from compounds pronounced by a German native speaker. The results of a perception test on these stimuli showed that the synthesis of the F0 contour was more effective than the synthesis of duration in correcting the lexical stress of the compounds (Bissiri 2006).

In general, language learners are often not able to recognize their wrong pronunciation without external help. Therefore they should receive appropriate feedback to identify and correct their own mistakes. Recast, "the teacher's correct restatement of a learner's incorrectly formed utter-

ance" (Nicholas, Lightbown & Spada 2001, p. 720), has shown to be an effective form of feedback for phonological mistakes (Lyster 1998). Recasts created by means of resynthesis of the learners' utterances should help them to better distinguish the difference between the right and the wrong pronunciation (Nagano & Ozawa 1990). According to this hypothesis some Computer Assisted Language Learning programs (CALL) offer as feedback the synthesis of the learners' utterances (Germain-Rutherford & Martin 2000; Hirose, Gendrin & Minematsu 2003; Hirose 2004; Martin 2004). Moreover the correction could be emphasized in the feedback (Chaudron 1977, "repetition with change and emphasis"), to help the learners identify their mistakes.

Consequently, feedback consisting in recorded utterances of a native speaker is not sufficient to make learners realize the wrong stress in their utterances. Our first hypothesis to be tested is that feedback in the learners' own voice instead, obtained with resynthesis, helps them better perceive the difference between the right and the wrong pronunciation (Tillmann & Pfitzinger 2004). Our second hypothesis is that emphasis of the correctly stressed syllable in the feedback helps the learners to more easily locate their errors.

In the next section we describe the method we used to test our hypotheses. In section 3 we report how we collected the data we used in the experiment. The synthesis procedure of the learners' speech is described in section 4. Section 5 explains how the pronunciation training was carried out. The perception test used to evaluate the training results is illustrated in section 6. Section 7 describes and discusses the results, and section 8 contains the conclusions.

2. Method

We recorded the story “Die Buttergeschichte”, a well-known standard German text, read by twelve Italian learners of German. The speakers read also a list of 32 words from the same text, including 16 compounds, whose first syllable has to be stressed. The same material was read twice by a German native speaker. First the native speaker read the utterances normally, then she emphasized the word stress on the first syllable of the compounds.

The Italian speakers of German were randomly assigned to a test and control group. The word stress of the compounds pronounced by the Italian speakers was checked for correctness by the first author and a German native speaker. The utterances by the speakers of the test group that contained compounds judged as not correctly stressed were prosodically corrected by means of synthesis. The synthesis was performed by copying from the utterance of the German native speaker the following prosodic parameters: local speech rate, fundamental frequency and intensity. Two versions were created: to obtain the first one we copied the prosodic parameters of the native speaker’s utterance having normal stress. For the second version we copied the parameters of the native speaker’s utterance with emphasized stress on the compounds.

During the training the Italian speakers heard two versions of each utterance: first, the one with the wrong stress they had produced during the first recording, then the prosodically corrected one. The speakers of the test group heard the resynthesized utterance in their own voice, and the control group heard the native speaker. After hearing this feedback they pronounced the utterance again, which was recorded. Then they repeated the exercise, but this time they received as feedback the version with emphasis. At the end of training we had two post-training versions of the same utterance for each speaker: the one after feedback without emphasis and the one after additional feedback with emphasis.

The correctness of stress in the pre- and post-training compounds was assessed in a perception test carried out by 37 German native speakers.

The goal was to statistically evaluate i) if feedback with synthesis in the learner’s own voice is more effective than feedback consisting in the recorded voice of a native speaker, ii) if feedback with emphasis is more effective than feedback without emphasis, iii) if feedback with emphasis and synthesis is the most effective.

3. Data collection

The recordings were made in an anechoic chamber at the Institute of Phonetics and Speech Communication of the University of Munich with the microphone Neumann TLM 103 at 48 kHz sampling rate and 16 Bit resolution. The German native speaker was a female speaker from Munich, 32 years old, and spoke with no dialect. The Italian speakers were ten female and two male speakers, between 20 and 52 years old, and living in Munich. The length of their stay in Germany was between 2 months and 32 years. Eight of them were advanced and four, including the two men, were beginning speakers of German.

The speakers read the text of the “Buttergeschichte” divided into 16 sentences, 14 of which contained 16 compounds in total to be stressed on the first syllable. Afterwards they read a list of 32 words from the “Buttergeschichte” as well, including the 16 compounds. The order of the words in the list was the same as in the story. The German native speaker was asked to read the utterances a second time putting emphasis on the word stress of the compounds in the sentences and on all the words of the list. She was suggested to imagine that a language learner would pronounce the words with stress on the wrong syllable, and that she was supposed to make him/her clear which syllable has to be stressed, just as a teacher would do in the classroom to help the student recognize the error.

The Italian speakers were randomly assigned to a test and a control group. To avoid that all four beginners, including the two male speakers, could end up in the same group, we randomly assigned one male and one female beginner to the test group and the remaining two to the control group.

The first author and a German native speaker examined the word stress of the compounds and of all the words from the list pronounced by each Italian speaker. The utterances by the speakers of the test group, whose words were judged as not correctly stressed, were submitted to synthesis to correct the mistake.

4. Synthesis

The synthesis was carried out with the tool *ProFIS*, *PROmpts For Information Systems* (Pfitzinger 2006).

4.1. Synthesis of the local speech rate

The term *local speech rate* indicates the speech rate that usually changes from one syllable to another and sometimes even within the same syllable (Pfitzinger 2001, p. 139).

The synthesis of the local speech rate was possible without a previous manual segmentation of the data. A Dynamic Time Warping algorithm (DTW) was used to align the phones in the signal of the Italian speaker with the phones in the signal of the German native speaker. For a future employment of this method in CALL programs, the synthesis without segmentation is very important. This way the learner could get the feedback in his own voice immediately after pronouncing the utterance, and he would be able to compare directly the correct and the wrong pronunciation with each other. For this experiment we did not adopt this approach for two reasons. First, the text material contained some quite long sentences. In these cases a preparation of the signal of the Italian speaker was necessary to manually align its pauses to those of the signal by the German native speaker. This happened especially when the two signals had a very different duration structure, for instance when the Italian speaker spoke very slowly and introduced some hesitations. Anyway, the DTW-algorithm worked well for single words and short sentences. The second reason was that we wanted to control the quality of each synthesized utterance, which we were going to use as feedback. It was important that the learner did not get bad quality utterances as feedback. For instance, some utterances were rejected

because in the signal of the Italian speaker the presence of creak or strong nasalization on the syllable to be stressed hindered a successful correction by means of synthesis. A high-quality synthesis was not possible for some utterances because the signals of the German and Italian speaker differed too much regarding duration or voice quality. We should mention that the Italian speakers had never heard a German speaker read this story before. This was a much more difficult situation than the usual training with a CALL program. Usually learners are asked to repeat a sentence after they heard it pronounced by a native speaker. If the learners in our study had read the utterances after hearing the pronunciation of a German reference speaker, the synthesis of their signals would probably have been easier, and this might be the situation if the technology we tested were implemented in a CALL program.

The duration of phones was modified using a PSOLA-algorithm so that the signal of the Italian speaker was made synchronous with the signal of the German native speaker.

4.2. Synthesis of the fundamental frequency

The analysis of the fundamental frequency was carried out with an island-driven AMDF-based algorithm (Average Differential Magnitude Function), which was developed by the second author. To keep the register of the Italian speaker in the synthesized version, the F0 contour extracted from the utterance of the German native speaker was divided by its mean value and multiplied by the mean value of the Italian speaker's F0 contour, before it was copied to it. By using this method and reducing the F0 frequency band in the detection, we could copy the F0 contour of the female German speaker to the signals of the male Italian speaker in the test group as well.

4.3. Synthesis of the intensity

Copying the intensity of the native speaker's utterance to the signal of the Italian speaker was necessary especially for emphasized stress because the intensity contour differences were in most cases perceptible. The amplitude envelope of the German speaker's signal was copied to the amplitude-normalized signal of the Italian speaker.

5. Pronunciation training

A week after the first recording, the Italian speakers came back for the training session. It is most unlikely that the speakers could have improved their pronunciation of compounds during that time. First of all, they had no idea what the training was going to be like. Secondly, most of them had been living in Germany for several years and had never noticed this mistake, which was fossilized in their pronunciation.

The training session for each of the Italian speakers was based on those of his/her utterances that contained wrongly stressed compounds. Therefore the material used in the training was different for every participant. Most speakers seemed to have more problems with the sentences: sometimes the wrong stress was found only in the sentences and not in the corresponding isolated words. In this case, even if the isolated compound was originally pronounced correctly, both the isolated compound and the utterance con-

taining it were included in the training of the speaker. This was done to draw the attention of the participant to the single word and to maximize the learning effect. If the Italian speakers wrongly stressed words other than the 16 compounds, these were also included in the training.

Training was carried out in the anechoic chamber as well. This time in front of the participant a loudspeaker was placed to present the prepared stimuli. The participants received a list with the words and the sentences, in which they had made a mistake. Since the training material was differently extensive for each participant, the duration of the training was different for everyone but never lasted longer than one hour.

The training consisted of two parts. In the first part the participants of the test group received the following instructions. They were told that they would hear two versions of each word and sentence, their original pronunciation and their pronunciation corrected according to the following prosodic characteristics: i) intonation, which is the melody, ii) the duration of the vowels and the consonants, which also means the "velocity of speech", iii) the intensity, which means the loudness of speech. They would hear each version twice. After that, the participants should pronounce the utterance twice imitating the correct pronunciation. The new pronunciation was then recorded. During the production of the words and sentences the participants should take care to correctly reproduce the prosodic characteristics described above. For each sentence coming after a word, they should pay particular attention to the same word contained in the sentence.

When the participants finished the first part of the training, the instructions for the second part were given: the participants would hear the two versions they heard before plus a third version, which consisted in a synthetically corrected utterance, in which the word stress was emphasized so that they could recognize it more easily and produce it correctly. The participants heard the three versions and pronounced the utterances again.

The training of the participants of the control group was structured in the same way except that they heard the correct utterance pronounced by the German native speaker instead of the corrected version in their own voice. In the first part of the training they were also told that they should take care to correctly reproduce the prosodic characteristics described above. All participants received a compensation for their work.

The use of synthesis in the pronunciation training had a motivating effect on the learners. The participants of the test group showed more interest in the training and were curious and surprised to hear the correct pronunciation in their own voice. At the end of the training, some even said they would be happy to come back in a week to do some more exercises of that kind. On the contrary, many participants of the control group were disappointed by the training and did not show the same interest as the test group.

6. Perceptual evaluation of training methods

To determine the effect of the training methods, the compounds pronounced before and after the training were evaluated in a perception test (see Fig. 1).

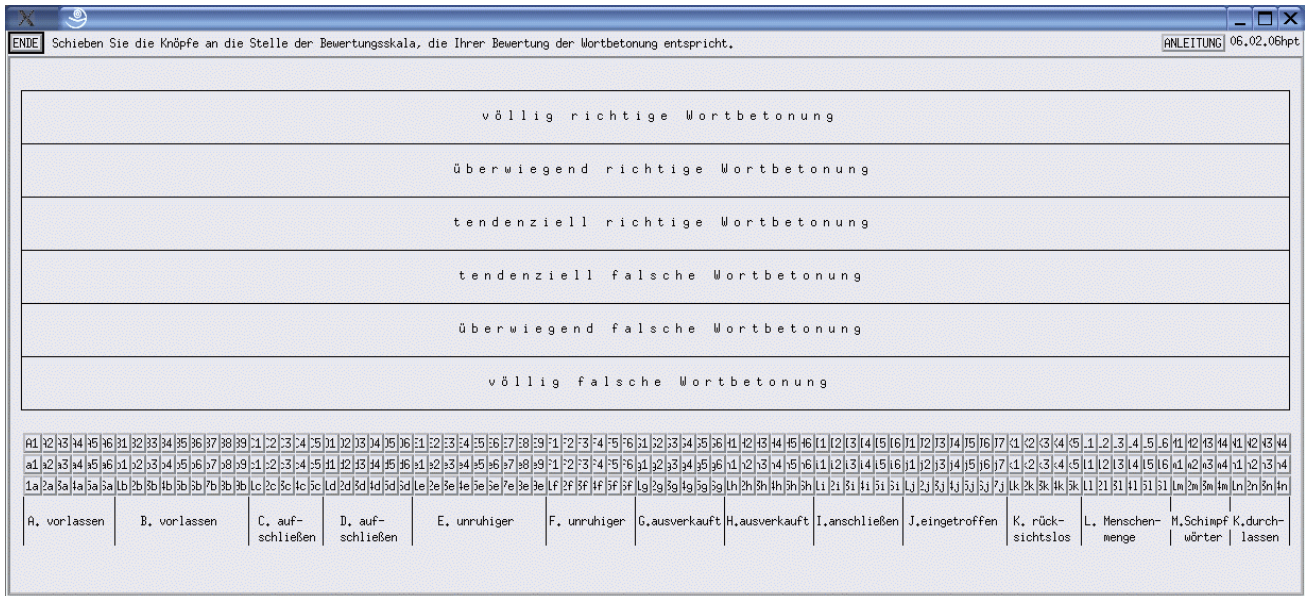


Figure 1: User interface for the word stress correctness assessment test.

Only the utterances which were originally judged to have compounds with wrong stress and their post-training correspondence were included in the test. Nevertheless, there were too many. So we decided to reduce the number of stimuli to 255. In order to choose the 85 pre-training utterances, which would correspond to 255 pre- and post-training utterances, we chose the most frequent wrongly stressed utterances that were used in the training. They were eight isolated compounds and six compounds embedded in sentences. This way we obtained for each word or sentence several stimuli from different Italian speakers, and the most frequent mistakes were also more relevant for the investigation.

The perception test was carried out by 37 German native speakers, between 21 and 53 years old. 25 of them were employees or students of the Institute of Phonetics and Speech Communication of the University of Munich.

The test was implemented with the tool CoDIT (Pfitzinger 2006). The stimuli were ordered in columns. The three buttons of each column, which had the same letter-number coding, corresponded to the pre-training version and the two post-training versions of the same utterance from the same speaker. Below the buttons, the evaluators could read the compounds whose stress they were going to evaluate (see Fig. 1). In each column the three versions were in random order. The order of the speakers was also randomized for each utterance. The evaluators should compare the three versions in each column with each other and drag them up into the upper area according to their judgement of the correctness of stress. The evaluators could choose between the following six areas to place the buttons: 1. completely correct word stress, 2. mostly correct word stress, 3. tending towards correct word stress, 4. tending towards wrong word stress, 5. mostly wrong word stress, 6. completely wrong word stress. The more correct the evaluators judged the stress of a compound, the further up they should locate the corresponding button. The evaluators could also express fine differences in the word stress, placing the buttons at different heights inside the same area.

7. Results and discussion

To evaluate the results, we converted the judgements of the native speakers to a range between 0 (meaning the worst possible judgement) to 100 (meaning the best possible judgement).

Table 1 shows the mean values of the three versions (pre-training, post-training without emphasis and post-training with emphasis). The judgements of the post-training versions produced after feedback with synthesis and/or emphasis yield the highest mean values.

Fig. 2 represents the improvement caused by feedback with no emphasis and feedback with emphasis for control and test group. The values on the y-axes are calculated by subtracting the judgement of the pre-training stimuli from the corresponding post-training stimuli. Feedback with emphasis produced in both control and test group a greater improvement than feedback without emphasis.

Since the data were not normally distributed, we used non-parametric tests for significance. To test if the judgements of the pre-training and of the two post-training versions are significantly different, the Wilcoxon test was employed. The results of the Wilcoxon rank sum test for unpaired samples show that there is no significant difference between the pre-training versions of test and control group ($p = 0.1099$). The level of proficiency in German language of both groups can be considered equivalent. The comparison of the post-training stimuli after feedback without emphasis between test and control group, instead, showed that feedback in the learner's own voice was more effective (Wilcoxon rank sum test, $p < 0.001$). The comparison was carried out by considering the improvement values as shown in Fig. 2.

	pre-training	no emphasis	emphasis
control group	37.35	55.21	62.70
test group	38.13	61.36	63.80

Table 1: Mean values of the three versions of the stimuli for control and test group.

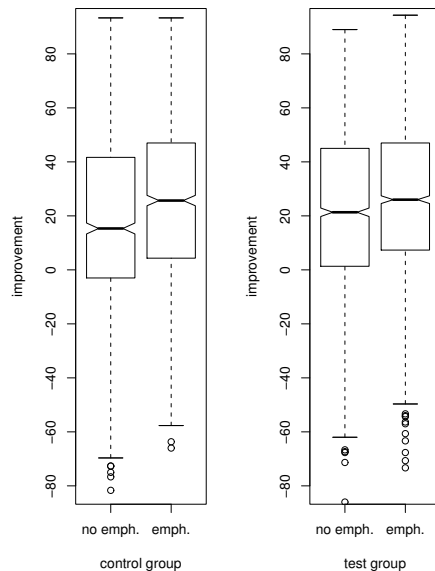


Figure 2: Improvement after feedback without emphasis and after feedback with emphasis for control and test group.

Applying the pairwise Wilcoxon signed rank test for paired samples to the three versions of the control group yielded significant differences between the pre-training and both post-training versions ($p < 0.001$). The post-training stimuli produced by the control group after feedback with emphasis were significantly better than those produced after feedback without emphasis ($p < 0.001$).

We also compared the three versions of the test group. Both the version after feedback with no emphasis and the one after feedback with emphasis received significantly better judgements than the pre-training version ($p < 0.001$). Between the two post-training versions there was an almost significant difference ($p = 0.052$). We had expected feedback in the learner's own voice with emphasis to be more effective than the same feedback without emphasis, and we would probably have had a higher significance with more data. Nevertheless many Italian speakers imitated the emphasized version, causing the evaluators to judge the word stress as correct, but assigning it a lower score than if the word had not been overstressed. Twelve of the 37 evaluators were asked about the presence of overstressed stimuli, and how they judged them. All twelve admitted they heard overstressed stimuli. Eight evaluators out of twelve said they judged the overstressed stimuli as right, because the speakers stressed the right syllable. However, they said they gave those stimuli a worse judgement than they would have done if they had been normally stressed.

Fig. 3 represents the medians of the three versions for the 52 utterances of the test group, located on the test interface. The medians are ordered from the lowest to the highest value according to the post-training version without emphasis. The medians of the post-training version after feedback with emphasis are located on the upper part of the area, which points out that they were judged as correctly stressed from the evaluators. However they are mostly judged worse than the stimuli after feedback without emphasis, indicating that they might have been "punished" by the evaluators possibly because of overstress. On the left part of Fig. 3 we can see 10 of the 52 stimuli after the feedback without emphasis that are located under the middle

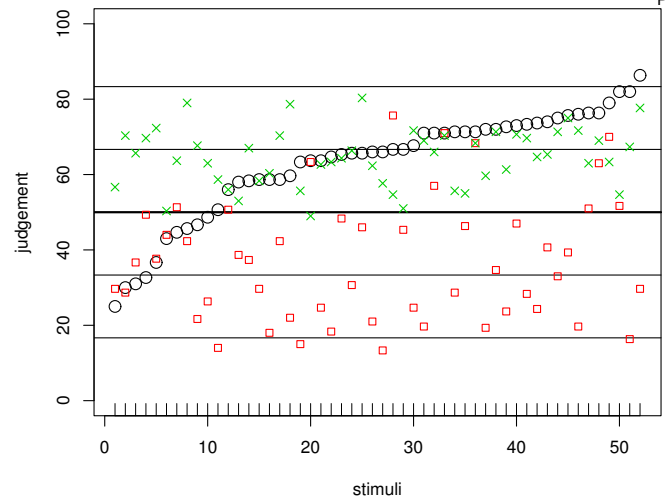


Figure 3: Median of the three versions of the test group. Red squares = pre-training without emphasis. Black circles = post-training without emphasis. Green crosses = post-training with emphasis.

line, and whose stress realisation was therefore judged as incorrect. The corresponding stimuli after feedback with emphasis were judged as more correct, and are located on the upper half of the area. In the case of these stimuli, the Italian speakers were possibly not able to pronounce the correct stress position after hearing the feedback without emphasis, and emphasis was necessary for them to recognize and reproduce the right position for stress.

The comparison of the improvement of test and control group by means of feedback with emphasis showed that feedback in the learner's own voice with emphasized stress was not more effective than the recorded native speaker with emphasized stress (Wilcoxon rank sum test for unpaired samples, $p = 0.5956$).

We had expected that feedback with emphasis in the learner's own voice would be more effective than feedback with emphasis in the recorded voice of the native speaker. A possible reason why this did not happen could be that the comparison between the incorrect and the correct version in the Italian speaker's own voice was sufficient for him to perceive the difference and that possibly adding emphasis did not bring much improvement. The results could have been also influenced by the tests design and by the fact that evaluators "punished" the overstressed versions. The evaluators had to compare three stimuli at a time: the pre-training and the two post-training versions. Since feedback without emphasis in the learner's voice was more effective than feedback without emphasis in the native speaker's voice, the evaluators may often have compared for the test group three versions, in which the post-training versions were both stressed on the right syllable. Therefore they possibly made a distinction between them by assigning the overstressed version a worse judgement. Instead, the post-training versions by the control group may have often contained one incorrect version, produced after feedback without emphasis, and one overstressed version, produced after feedback with emphasis. Comparing the overstressed version with the incorrect version the evaluators possibly did not punish it as much as they would have done if they had had to compare it with a correct version.

8. Conclusions

Language learners should be put in a position to understand the difference between the right and the wrong stress pronunciation, in order to learn to speak a foreign language correctly. This applies to Italian speakers, who are possibly not able to notice the correct word stress of German compounds without receiving appropriate indications. In this case, according to our hypothesis, appropriate forms of feedback could be i) feedback in the learner's own voice and ii) feedback with emphasized word stress.

The main outcome of this study is that utterances in the learner's own voice and/or with emphasized word stress are a more effective form of feedback for stress pronunciation training than pre-recorded reference utterances spoken by a German native speaker.

Since synthesis of local speech rate can be carried out without manual segmentation, such a technology could be used to create more effective CALL programs. In our study, feedback with synthesis showed a motivating effect on the learners. Results show that synthesized utterances in the learner's own voice without emphasis are more effective than utterances of a native speaker without emphasis. Nevertheless synthesis combined with emphasis was as good as feedback in a reference speaker's voice with just emphasis.

Possible reasons for this result could be i) combining both parameters emphasis and synthesis does not bring much improvement, since feedback based on one of them might often be sufficient for the learner to recognize the difference between the right and the wrong pronunciation, and ii) evaluators might have "punished" overstressed stimuli especially if coming from the test group, since they possibly had to compare them with another correct stressed stimulus more often than in the control group.

The difference between the effect of feedback in the learner's own voice with emphasis and without emphasis was almost significant. With more data we could have achieved higher significance. Nevertheless, the fact that overstressed stimuli were judged as less correct might be the cause for this result.

Finally, it turned out very clearly that utterances produced by a native speaker with emphasis are more effective than utterances of a native speaker without emphasis.

We did not find enough Italian participants to build four groups: 1. feedback with the traditional method, 2. feedback with synthesis, 3. feedback with emphasis, and 4. feedback with emphasis and synthesis. In this case the evaluators would have compared two instead of three stimuli at a time, and the post-training version would have only been compared with the original pre-training version thus avoiding the comparison between two post-training stimuli.

If we had had four groups of Italian participants, we could have excluded any carry-over effect between the two parts of the training. Still, in the second part of the training, the Italian learners were provided with a new information by means of the emphasized stress, which they did not get the first time. Hearing the utterances without emphasized stress several times would not have had the same effect, because the learners were not put in a position to notice the correct position of stress.

References

- Bertinetto, P. M. (1980). Perception of stress by Italian speakers. *Journal of Phonetics* 8, 385–395.
- Bertinetto, P. M. (1981). *Strutture prosodiche dell'italiano. Accento, quantità, sillaba, giuntura, fondamentali metrici*. Firenze: Accademia della Crusca.
- Bissiri, M. P. (2006). Einfluss von Intonation und lokaler Sprechgeschwindigkeit auf die Betonung von durch L2-Sprecher ausgesprochenen Komposita. In C. D. Meola, L. Gaeta, A. Hornung, and L. Rega (Eds.), *Akten der zweiten Tagung Deutsche Sprachwissenschaft in Italien*. Roma: (Roma, 9.–11.2.2006). Istituto Italiano di Studi Germanici, (accepted).
- Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learner's errors. *Language Learning* 27, 29–46.
- Germain-Rutherford, A. and P. Martin (2000). Présentation d'un logiciel de visualisation pour l'apprentissage de l'oral en langue seconde. *Apprentissage des Langues et Systèmes d'Information et de Communication (ALSIC)* 3(1), 61–76.
- Hirose, K. (2004). Accent type recognition of Japanese using perceived mora pitch values and its use for pronunciation training system. In *Proc. of the Int. Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages (TAL)*, Beijing, pp. 77–80.
- Hirose, K., F. Gendrin, and N. Minematsu (2003). A pronunciation training system for Japanese lexical accents with corrective feedback in the learner's own voice. In *Proc. of EUROSPEECH, Vol. 4*, pp. 3149–3152.
- Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning* 48, 183–218.
- Martin, P. (2004). Win Pitch LTL II, a multimodal pronunciation software. In *Proc. of InSTIL/ICALL2004 Symposium on Computer Assisted Language Learning*, Venezia, pp. 177–182.
- Nagano, K. and K. Ozawa (1990). English speech training using voice conversion. In *Proc. of Int. Conf. on Spoken Language Processing (ICSLP)*, Kobe, pp. 1169–1172.
- Nicholas, H., P. M. Lightbown, and N. Spada (2001). Recasts as feedback to language learners. *Language Learning* 51, 719–758.
- Pfützing, H. R. (2001). Phonetische Analyse der Sprechgeschwindigkeit. In *Forschungsberichte (FIPKM)* 38, pp. 117–264. Inst. für Phonetik und Sprachliche Kommunikation der Univ. München.
- Pfützing, H. R. (2006). ProFIS (PROMpts For Information Systems), CoDIT (COMbined Discrimination/Identification Task). <http://www.phonetik.uni-muenchen.de/~hpt/#Software>, 05.09.2006.
- Tillmann, H. G. and H. R. Pfützing (2004). The development of an advanced SLP-based system for the individual learning and fast training of speaking skills in a new foreign language. In *Proc. of InSTIL/ICALL2004 Symposium on Computer Assisted Language Learning*, Venezia, pp. 17–20.