Lexical stress perception of German morphologically complex words by Italian speakers

Maria Paola Bissiri [a/b] Hartmut R. Pfitzinger [c/a] Hans G. Tillmann [a]

[a] Institut für Phonetik und Sprachverarbeitung, University of Munich, Germany
[b] Dipartimento di Scienze dei Linguaggi, University of Sassari, Italy
[c] Institut für Phonetik und digitale Sprachverarbeitung, University of Kiel, Germany

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

Abstract

Previous research has shown that Italian speakers have difficulties in learning the correct lexical stress of German morphologically complex words. To test if this difficulty is of perceptual origin, we employed the method by Dupoux et. al 2001 [7], consisting in a short-term memory sequence repetition task. We tested the participants with a stress contrast, a phoneme contrast, which constituted the baseline, and a stress+quantity contrast, which should correspond to an Italian stress pattern, in which duration plays a role. We also tested a control group of German subjects. As expected, the Italian subjects performed worse on stress contrast than on phoneme or stress+quantity contrasts. The German subjects performed equally well on all contrasts, except for the stress contrast of words with long penultimate syllable, in which their error rate was comparable to that of Italian subjects.

1. Introduction

Italian speakers of German often have difficulties in learning the correct lexical stress of German morphologically complex words, such as compounds or separable prefix verbs, whose first component has to be stressed, since they tend to stress the second component instead. According to our hypothesis, Italian speakers make these errors because they have difficulties in perceiving the actual stress position. The hypothesis is supported by previous investigations carried out by the first author [4], which show that Italian proficient speakers of German, living in Germany since even ten years or more, never noticed this mistake and make it systematically. This is an aspect of pronunciation that is not difficult to reproduce for Italian speakers, provided that they perceive it. Moreover, an investigation consisting in lexical stress training for Italian speakers [3] showed that stimuli with emphasized stress or prosodically corrected synthetic stimuli in the learner's own voice are a more effective form of feedback than normally stressed or natural stimuli. Emphasis is helpful, since Italian speakers need a cue to locate the stress position. The comparison of their wrong pronunciation with the corrected synthetic stimuli in their own voice reduces the phonetic variability and helps them to better recognize the difference between correct and incorrect word stress.

In a previous investigation [4], the stress of wrongly pronounced morphologically complex words by Italian speakers was corrected by copying F0 and segment duration from the speech of a German native speaker. The results of a perception test showed that the correction of F0 was more effective than the correction of segment duration, which means that the wrongly stressed compounds by the Italian speakers deviated from the correct stress pronunciation mostly because of their F0 contour. If we now consider that in Italian duration has been demonstrated to be the strongest stress correlate [2], we suppose that, since Italian speakers rely on this parameter, they could have more difficulties in associating word stress with higher F0, especially if the unstressed syllables have long duration. In fact, in a German compound in which the first component carries the primary stress, the quantity pattern of the second component does not change, so that the stressed syllable of the second component keeps its long duration [8, p.188].

Our aim is to test if the Italian speakers' difficulty in learning the correct stress position in German morphologically complex words depends on their native-language-specific phonological representation of stress, which causes them to assimilate German minimal pairs such as 'umfahren and um'fahren to the same stress pattern. Italian learners of German might be able to solve without problems a task consisting in the simple discrimination or identification of words belonging to German minimal stress pairs, since stress is expressed by several acoustic parameters, which they can auditorily perceive. However this kind of test would not reveal how these words are processed at a more abstract level. For this reason, to test our hypothesis we employed a method by Dupoux et. al 2001 [7], involving memory load, phonetic variability, and limited time, so that the test persons have less possibilities to exploit acoustic strategies to solve the task.

Dupoux et. al 2001 [7] tested what they called "stress deafness" by French speakers. Since stress is non-contrastive in French, French native speakers should have difficulties to categorize words that differ just in stress position. The test for the French native speakers consisted in a short-term memory sequence repetition task. The participants should first learn to associate two non-words of a stress minimal pair to the computer keys *1* and *2*. Then they heard non-word sequences of two till six elements and should reproduce the sequences typing the keys in the correct order. The method has also been used to test speakers of other non-stress languages: Finnish, Hungarian, and Polish [9].

In our study we apply this method to speakers of the stress language Italian, which presents several minimal stress pairs. Since stress is language-specific, difficulties can arise in the perception of stress in a foreign language, if in this language the acoustic parameters are exploited by stress in a different way than in the native language. Despite the considerable number of stress minimal pairs, according to Bertinetto 1976 [1] Italian does not have minimal pairs characterized by an exchange of primary and secondary stress, like 'über͵setzen and ͵über'setzen in German. For Bertinetto secondary stress is a phenomenon

absent in languages that do not have composition capabilities, such as roman languages, and is present in germanic languages instead. The reason should be that in germanic languages the stress unit is not the word, like in Italian, but the morpheme [1, p.199-200].

2. Method

The method in our study is based on Experiment 4 by Dupoux et al. [7] with some differences. While they varied the pitch of the stimuli by means of resynthesis, we used just natural stimuli since we did not want to add any disturbance caused by unnatural pitch contours. The required phonetic variability was provided by stimuli spoken by a female and a male speaker. The participants were tested with i) a phoneme contrast, ii) a German stress contrast, and iii) a stress+quantity contrast, which should be more similar to the stress contrast in Italian, where stress is "supported" by vowel duration. The phoneme contrast was the same for every participant. The stimuli for the other two contrasts were randomly selected among 10 different word pairs. The same test was performed by German subjects as control group.

According to our hypothesis, the Italian subjects should perform worse on stress contrast than on phoneme or stress+quantity contrast. The German subjects should perform comparably well on these three contrasts. Since we used German words and pseudo-words, the German were advantaged over the Italian subjects. However, before using non-words unknown to both participant groups, we wanted to perform the test using German words in order to see if the hypothesized effect would arise at all.

2.1. Stimuli

To create the stimuli for the test we searched the CELEX database [5] for German three-syllabic words with a morpheme boundary between the first and the second syllable and the vowel /a/ in the first two syllables. We randomly selected five words with a long /a/ and five with a short /a/ in the penultimate syllable. All words should be produced in three versions: i) correctly stressed on the first syllable, ii) stressed on the penultimate syllable, iii) different for the two word groups and with a quantity change in the penultimate syllable: a) the words with a long /a/ should be read with a short /a/ in the penultimate syllable, and b) the words with a short /a/ should be stressed on the penultimate syllable and pronounced with a long /a/. This way two word pairs for each word could be created: one stress contrast and one stress+quantity contrast (s. Table 1). The quantity should be expressed by orthography. We did not choose existing German stress minimal pairs as stimuli, since there are none with an /a/ in the first two syllables. With other vowel qualities than /a/ the German quality change would be a cue.

For the phoneme contrast a phoneme minimal pair was also randomly chosen from the database. The words in the phoneme contrast should be stressed on the penultimate syllable, in order not to present to the Italian participants an unusual stress pattern in the baseline contrast.

The corpus was read by two German native speakers, trained phoneticians: a male speaker aged 30 and a female speaker aged 22. Their knowledge of Italian was none in one case, very sparse in the other. The recordings were made in an anechoic chamber at the Institute of Phonetics of the University of Munich with the microphone Neumann TLM 103 at 48 kHz sampling rate and 16 Bit resolution. The text was pre-

Table 1: *Stimuli represented according to stress and quantity of /a/ in the penultimate syllable or to the preceding consonant.*

Stimuli for the stress and for the stress+quantity contrast		
long and unstr.	long and stressed	short and unstr.
1 'Stadtplanung	*Stadt'planung	*'Stadtplannung
2 'Anlage	*An'lage	*'Anlagge
3 'abstrahlen	*ab'strahlen	*'abstrallen
4 'abnabeln	*ab'nabeln	*'abnabbeln
5 'Astgabel	*Ast'gabel	*'Astgabbel
short and unstr.	short and stressed	long and stressed
6 'absacken	*ab'sacken	*ab'saaken
7 'abflachen	*ab'flachen	*ab'flaachen
8 'anklammern	*an'klammern	*an'klahmern
9 'abzwacken	*ab'zwacken	*ab'zwaacken
10 'anstarren	*an'starren	*an'stahren
Stimuli for the phoneme contrast		
alveolar lateral-approxim.	bilabial nasal	
11 *ab'lagern	*ab'magern	

sented on a screen, and the recording was carried out with the tool SpeechRecorder [6]. To avoid contrast accent, each word was read three times, and only the third realisation was used in the perception test. The speakers read the whole material five times. Since we needed six tokens of each word, three realisations out of five were chosen for each speaker.

The test stimuli consisted in 10 words × 3 versions × 6 tokens = 180 stimuli for the stress and the stress+quantity contrasts and 2 words × 6 tokens = 12 words for the phoneme contrast. All stimuli were segmented and subjected to acoustic measurements of syllable and vowel duration.

The changes in vowel quantity influenced syllable and vowel duration. We compared the equally stressed versions with quantity difference. The penultimate syllables of the versions pronounced shortening the vowel are on average 79.1 ms shorter than the original versions with long vowel. The penultimate syllable of the versions pronounced lengthening the vowels are on average 106.8 ms longer than the versions with the short vowel. The same comparison for the vowel /a/ in the second syllable shows that shortened vowels are on average 81 ms shorter and lengthened vowels are on average 102.3 ms longer than their counterparts. In all cases the differences are highly significant ($p < 0.001$, t-test with pairwise comparisons).

Stress changes did not always influence syllable and vowel duration. We compared the first and then the second syllables and vowels of the words in the stress contrast. The first syllables in the words with long penultimate are on average 46.8 ms shorter if unstressed. The first syllables in the words with short penultimate are on average 33.7 ms shorter if unstressed. The difference is in both cases significant ($p = 0.002$ and $p = 0.0419$). The differences between the second syllables are not significant. The same comparison carried out with the vowels shows that the unstressed vowels in the first syllable are on average 10.3 ms shorter in the words with long vowel in the penultimate and 15.4 ms shorter in the words with short vowel in the penultimate. The difference is almost significant in the first case ($p = 0.052$) and highly significant in the second case ($p < 0.001$). The long unstressed vowels in the penultimate are on average 37.5 ms shorter than their stressed counterpart. The short unstressed vowels in the penultimate are on average only 3 ms shorter. In the first case the difference is highly significant ($p < 0.001$), while in the second case it is not significant.

2.2. Subjects

The test persons were 23 Italian native speakers, between 20 and 51 years old. They spoke no German. Only one participant reported having very sparse knowledge of German, but he declared he had never heard the words in the test before. The participants of the control group were 6 German native speakers, between 14 and 49 years old. All Italian and German participants were phonetically naive subjects.

2.3. Perception test

The perception test was implemented in Java by the first author.

All participants were first tested with the phoneme contrast. Then they were tested with one stress contrast and one stress+quantity contrast. The order of presentation of these two contrasts was random for each participant.

The procedure was as in [7]. The participants first learned to associate the two words in the contrast with the computer keys *1* and *2*. They first listened to all six tokens for each word. Then they could press the keys to hear one token at a time as many times as they wanted. When they felt ready, they did a test to check if they had learned the word-key association. They listened to single words and tipped the key correspondent to the word they heard. If the answer was correct, the green text "OK!" was displayed for 800 ms on the screen. Otherwise the red text "ERRORE!" was displayed. In order to pass the test, participants should give five right answers in a row.
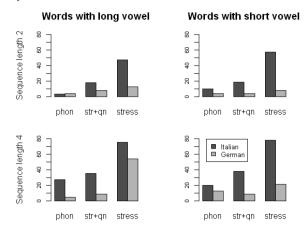
After that, the subjects listened to eight word sequences of two and eight of four elements and had to press the computer keys correspondent to the words in the sequence in the right order (example: ˈAnlage, ˈAnlage, *Anˈlage, ˈAnlage — 1-1-2-1). In this phase the test persons received no feedback. We did not present six word sequences in the test, because in a pilot test, carried out with one Italian and one German subject, the participants reported that it was impossible for them to remember the sequences of six elements of any contrast. The four possible two-word sequences were proposed twice. Out of the 16 possible four-word sequences the same as in [7] were chosen. The sequences were presented in random order. The tokens were randomly instantiated, so that half of the tokens in each sequence were spoken by the female and half by the male speaker, and each token could not appear twice in the same sequence.

The words in each sequence were separated by a pause of only 20 ms, so that the participants had less chances to translate the words into numbers while listening. A pause of 1500 ms separated each sequence from the next one. At the end of each sequence the participants heard the word "okay" pronounced by a male English native speaker, in order to prevent echoic memory (s. [7] and references cited there). The participants could not start tipping their response until they heard this word.

The experiment lasted about 15-20 minutes.

As Dupoux et al. [7] we coded only the 100% correctly reproduced sequences as correct and the 100% wrong responses as reversals. Dupoux et al. [7] rejected participants who had more reversals than correct responses either in the phoneme or in the stress contrast, since they possibly confused the association between words and computer keys. For our criterion, in order to be rejected i) a participant should have more reversals than correct responses in the two- or in the four-word sequences in at least one of the contrasts, ii) the number of reversals should be greater or equal to the critical value of the most powerful one-tailed binomial test with α-level lower than 0.05, and therefore at least 5 (α-level = 0.0273) in the two-word and at least 3 (α-level = 0.0108) in the four-word sequences. These thresh-

Figure 1: *Percent error with the three contrasts for the two- and four-word sequences. Three German and eleven Italian subjects were tested with contrasts created from German words with long /a/ and the rest three German and twelve Italian subjects with contrasts from words with short /a/ in the penultimate syllable.*

olds were calculated by means of the binomial quantile function, considering the probability to produce reversals by chance (1/4 for the two-word and 1/16 for the four-word sequences) and the number of trials (8). According to these criteria four Italian participants out of 27 were rejected.

3. Results and discussion

After finishing the test the Italian participants were asked which contrast was the most difficult to distinguish and which was the difference between the words in each pair. According to 12 of 18 subjects who had an answer to the question, the difference between the words of the stress+quantity contrast consisted in stress. This supports our hypothesis that the stress+quantity contrast would be assimilated to a stress contrast by Italians. The word pair corresponding to the actual stress contrast was the most difficult for the Italian participants to distinguish. Several declared they heard almost no difference between the words. However all Italian participants were able to produce, even after a long sequence of trials, the five correct responses in a row necessary to access the test phase. Therefore, they were able to distinguish the stimuli acoustically.

The percent of errors of the Italian participants is consistently higher in the stress contrast than in the phoneme and in the stress+quantity contrasts (s. Fig. 1 and Table 2). In the latter contrast Italian participants made more mistakes than in the for-

Table 2: *Percent error for the Italian and German participants as in Fig. 1.*

Contr. from	words with long /a/			words with short /a/		
Seq. length	2	4	mean	2	4	mean
Italian						
phoneme	3.4	27.3	15.3	10.4	19.8	15.1
stress+qnt.	18.2	35.2	26.7	18.8	37.5	28.1
stress	47.7	75.0	61.4	57.3	78.1	67.7
German						
phoneme	4.2	4.2	4.2	4.2	12.5	8.3
stress+qnt.	8.3	8.3	8.3	4.2	8.3	6.2
stress	12.5	54.2	33.3	8.3	20.8	14.6

mer contrast. The behaviour is comparable in the two groups, the one tested with stimuli originating from German words with a long penultimate and the other tested with stimuli from German words with a short penultimate syllable.

Instead, the German participants behave differently in the two groups. In the group tested with contrasts originated from words with a short vowel in the penultimate syllable they match the expectations and show comparable amount of errors in the three contrasts. The group of German participants tested with the stress contrasts with a long vowel in the penultimate had more trouble memorizing the four word sequences than the group tested with stress contrasts with a short vowel in the penultimate. Since there were only three German participants per group, this effect might be casual. However the same participants had no problems with the phoneme and stress+quantity contrasts.

In order to account for the repeated measures, we fitted the number of wrong responses with a generalized linear mixed model with multivariate normal random effects using Penalized Quasi-Likelihood, poisson error distribution, and logarithmic link function.

We found no significant difference between the performance of the Italian and of the German participants of the two groups (tested with word pairs with long or short vowel in the penultimate) on phoneme baseline contrast both for the two-word and the four-word sequences. As hypothesized, the Italian subjects performed significantly worse on stress contrast than on phoneme and stress+quantity contrast. (s. Table 3).

Table 3: *P-values resulting from performance comparisons.*

Contrast from	words w. long /a/		words w. short /a/	
Sequence length	2	4	2	4
Italian				
phon. vs stress+qnt.	0.0307	n.s.	n.s.	0.0187
stress vs stress+qnt.	0.0095	0.0011	0.0086	0.0005
stress vs phon.	0.0010	0.0002	0.0021	0.0000
German				
phon. vs stress+qnt.	n.s.	n.s.	n.s.	n.s.
stress vs stress+qnt.	n.s.	0.0273	n.s.	n.s.
stress vs phon.	n.s.	0.0270	n.s.	n.s.

We expected that the German subjects would perform equally well on the three contrasts. However, on four-word sequences of the stress contrast with long /a/ in the penultimate syllable they performed significantly worse than on phoneme contrast and on correspondent stress+quantity contrast, and they were not better than the Italian subjects (p = 0.0996).

Nevertheless considering both word groups and sequences, the German subjects were significantly better than the Italian subjects in the stress contrast (p<0.001) and almost significantly better in the stress+quantity contrast (p = 0.0583).

The performance of 12 Italian subjects on two-word sequences and 17 Italian subjects on four-word sequences of the stress contrast was not significantly better than chance (p-value>0.05, binomial test). However, the performance of the whole sample on two-word and four-word sequences of the stress contrast was in both cases better than chance (p<0.001).

4. Conclusions

Italian speakers of Geman have difficulties in learning the correct lexical stress of German morphologically complex words, because their language-specific phonological representation of stress causes them to assimilate to the same stress pattern differently stressed words of this kind. As hypothesized, Italian subjects presented with a short-term memory sequence repetition task performed worse on stress contrast than on phoneme or stress+quantity contrast. The latter contrast was interpreted by most participants as a stress contrast since stress was "supported" by a vowel duration change, and was therefore more similar to a stress contrast in Italian. Duration measurements of the stimuli showed that in some cases the stressed vowels or syllables and their unstressed counterparts did not differ in duration, and this possibly made the stress contrast difficult to differentiate for the Italian subjects.

The German test persons performed equally well on the three contrasts except, surprisingly, on stress contrast of words with a long penultimate syllable. However, this could be due to the small number of German participants. We intend to test a larger number of German and Italian participants to get more reliable results. Moreover we plan to carry out the same experiment with non-words unknown to both German and Italian speakers, in order to test both populations on an equal base.

5. References

[1] Bertinetto, P.M., 1976. L'accento secondario nella fonologia italiana. Analisi teorica e sperimentale. *Studi di fonetica e fonologia. Atti del convegno internazionale di studi*, R. Simone, U. Vignuzzi, and G. Ruggiero (eds.). Roma: Bulzoni, 189-235.

[2] Bertinetto, P.M., 1981. *Strutture prosodiche dell'italiano. Accento, quantità, sillaba, giuntura, fondamenti metrici.* Firenze: Accademia della Crusca.

[3] Bissiri, M.P.; Pfitzinger, H.R.; Tillmann, H.G., 2006. Lexical stress training of German compounds for Italian speakers by means of resynthesis and emphasis. *Proc. of the 11th Australasian Int. Conf. on Speech Science and Technology (SST 2006)*. Auckland, 24-29.

[4] Bissiri, M.P., 2007. Einfluss von Intonation und lokaler Sprechgeschwindigkeit auf die Betonung von durch L2-Sprecher ausgesprochenen Komposita. *Perspektiven Zwei. Akten der 2. Tagung Deutsche Sprachwissenschaft in Italien*, C. Di Meola, L. Gaeta, A. Hornung, and L. Rega (eds.). Roma: Ist. Italiano di Studi Germanici, 353-364.

[5] CELEX, 2004. *The CELEX lexical database.* Max Planck Institute for Psycholinguistics. Nijmegen. http://www.ru.nl/celex/, 16.08.2004.

[6] Draxler, C.; Jänsch, K., 2004. SpeechRecorder a Universal Platform Independent Multi-Channel Audio Recording Software. *Proc. of the 4th International Conference on Language Resources and Evaluation (LREC)*. Lisboa.

[7] Dupoux, E.; Peperkamp, S.; Sebastián-Gallés, N., 2001. A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America*, 110(3), 1606-1618.

[8] Kohler, K.J., 1995. *Einführung in die Phonetik des Deutschen.* Berlin: Erich Schmidt Verlag, 2nd edition.

[9] Peperkamp, S.; Dupoux, E., 2002. A typological study of stress 'deafness'. *Laboratory Phonology*, 7, 203-240.