A PROBLEM OF PROSODIC TRANSFER IN THE PERCEPTION OF GERMAN LEARNERS OF JAPANESE AND JAPANESE LEARNERS OF GERMAN

Mayako Niikura[a], *Tsutomu Sugawara*[b] *& Ursula Hirschfeld*[c]

[a]Department of German Language and Studies, Sophia University, Japan;
[b]Phonetic Laboratory, Sophia University, Japan;
[c]Martin-Luther Universität Halle-Wittenberg, Germany
m-niikur@sophia.ac.jp, sugawara@sophia.ac.jp, ursula.hirschfeld@sprechwiss.uni-halle.de

ABSTRACT

This study investigates F0, intensity and duration as acoustic parameters for the perception changing their relative weight, according to the perceptual development. Four groups – native Japanese who have studied German for one year at the Department of German, native Germans who don't understand Japanese, Japanese students who had lived in Germany for more than one year, and German students who had lived in Japan for about one year – were given a perception test of manipulated 2-syllable words. The results indicate the following: 1) Japanese students employ F0 as a perception cue. 2) German students employ all of the acoustic parameters as perception cues: F0, intensity and duration. 3) Japanese may also employ intensity as a cue at the similar level to Germans depending on the length of time they have learnt German and lived in Germany. 4) Germans learning Japanese and living in Japan tend to reduce the use of intensity as a cue.

Keywords: L2, speech perception, prosodic transfer, Japanese, German

1. INTRODUCTION

This study examines how the acoustic features which determine lexical stress in German change in their use during the process of German and Japanese learners studying the other language (Japanese and German). I.e., how the perceptual weight of the three parameters changes. It also examines how the use of the acoustic cues changes with the length time spent learning the L2 and living in the country where the L2 is spoken. It looks at which acoustic cues Japanese learners of German use to perceive lexical stress in German and how they change their perception as they learn German. Similarly, how Germans perceive lexical stress and how their perception changes as they learn Japanese. This type of bidirectional transfer has rarely been examined, although there is a study of the production of Japanese by Ueyama [12]. But by taking into consideration the bidirectional transfer we can compare perception at the prosodic level, and find how the prosodic elements which are not found in both languages influence each other.

It is accepted that the stress patterns of L1 influence the learning of stress in L2. The L1 background of non-native learners influences perception of L2 prominence and they transfer the perceptual strategies of L1 already learned to L2 [2, 6]. The lexical stress patterns of Japanese and German are different. German is a stress-accent / stress-timed language, while Japanese is a pitch-accent / non-stress accent [1] / mora-timed language. In Japanese, F0 ordinarily falls from an accented mora to following mora, but intensity hardly varies [8]. In German lexical accent is manifested in the acoustic parameters of fundamental frequency, duration, and intensity.

According to Kohler [11], the most important parameter is pitch; intensity and duration are secondary. Dogil & Williams [5], on the other hand, claim that duration is the most closely linked to lexical stress. Jessen, et al. [10], however, say vowel quality is the most related to word stress, and that the duration of tense vowels also influences stress perception. Dogil [4] concludes that F0 is related rather more to sentence stress than to word stress and intensity is linked to word stress, but is not so influential as the other parameters (F0 and duration).

2. EXPERIMENT

2.1. Participants

There are four groups, of which one is a control group (NG), comprising two Germans aged 26 and 43, who have not studied Japanese. They are

phonetic trained and are involved in phonetic research. They provide a baseline with which to compare the prosodic change of the following 3 groups: J(g) = 10 Japanese students (Sophia University) who have been learning German for more than four years and have lived in Germany for more than 1 year. Their average age is 23 (22 – 24). G(j) = 6 German exchange students from Berlin, Bonn, Cologne and Trier who have been learning Japanese for 2 – 4 years in Germany and who have lived in Japan for an average of eight months (1 – 11 months) and are learning Japanese at Sophia University. Their average age is 23 (22 – 26). NJ = 20 Japanese freshmen in the German Department of Sophia University. They have been learning German for 1 year and have never lived in Germany. Their average age is 19 (19 – 20). All of them listened to the randomized stimuli through headphone. The CALL-Room was used to test the 20 NJs as a group and a very quiet room was used to test the others subjects individually.

2.2. Procedure

We used acoustic manipulation to investigate the stress perception of lexical tokens of "mama", a word chosen for the following reason: Unlike other vowel pairs in German, the quality of the two vowel pairs /a/-/a:/ and /ɛ/-/ɛ:/, do not vary with length, making them both good candidates for duration manipulation. Japanese, however, has /a/, but no /ɛ/ in its phonological inventory, which precludes the latter. German "mama" is acceptable with stress on either the first or the second syllable, a sine qua non for the judgment of stress location.

The speakers who recorded German "mama" are all professors in the Department of German at Sophia University, who are accustomed to reading aloud. They were asked to record the first and second syllables of the target with and without stress three times into PCM recorder (Sony PCM-DJ). The sampling rate was 44.1 kHz, with 16 bit quantization. The clearest syllables were used for manipulation using PRAAT [3]. Syllable 1 of "mama" was modified to give five levels of F0 and intensity and five durations. The values were set such that level 1 was considerably higher (50Hz, 9dB, 50% longer) than the corresponding value for syllable 2. Level 2 was slightly higher (25Hz, 4.5dB, 25% longer). Level three had equal values to syllable 2. Level 4 was slightly lower (-25Hz, -4.5dB, 25% shorter) and level 5 was considerably lower (-50Hz, -9dB, 50% shorter). Table 1 summarizes the manipulation.

Table 1: Manipulations of F0, intensity and duration carried out on the base stimulus "mama".

	level 1	level 2	level 3	level 4	level 5
syll. ratios	syll1>syll2	syll1>syll2	syll1=syll2	syll1<syll2	syll1<syll2
F0	+50 Hz	+25	0	-25	-50
intensity	+9 dB	+4.5	1	-4.5	-9
duration	1.5 multiplied	1.25	1	0.75	0.5

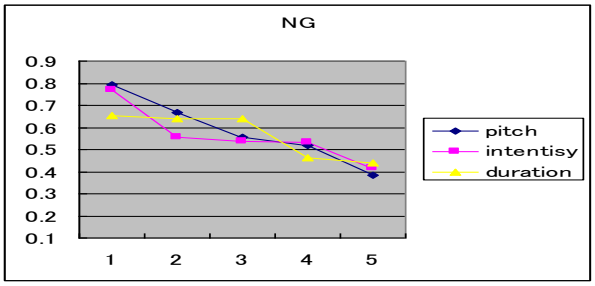
Each level of each parameter was combined with the other parameter values, giving 125 tokens (5 x 5 x 5) with all combination of F0, intensity and duration. These were presented to the subjects three times in random order, eliciting 375 responses per subject.

2.3. Results and discussion

Altogether, 14250 responses (38 subjects x 375 tokens) were analysed. Only the tokens judged to have stress on the 1st syllable are used for analysis, allowing the average of ratio of stress on the 1st syllable to be calculated. Analysis of variance (ANOVA) is used for statistical evaluation to examine the contribution of F0, intensity and duration to the perception of lexical stress by each subject group.

Figures 1 and 2 present the results for the NG and NJ group, respectively. The X axis shows the degree of manipulation where 1 to 5 correspond to levels 1 to 5 in Table 1, The Y axis shows the percentage of subjects who judge that the stress falls on the 1st syllable.

Figure 1: Results for 3 parameters in the NG group.

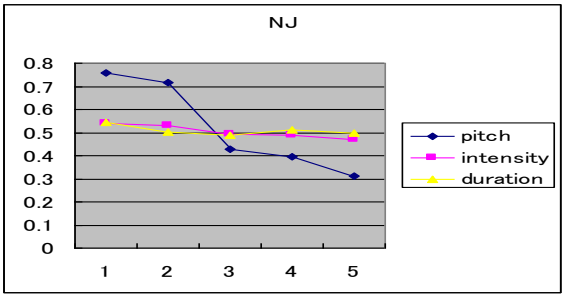


The German informants in the control group NG show the decline of pitch-dependent stress judgments from 79% for a level 1 F0 value to 38% for a level 5 F0 value (note that this is independent of the varying values of the other parameters in these stimuli). Intensity-dependent judgments dropped from 76% for level 1 intensity values to 41% for level 5 intensity value and duration-dependent judgments dropped from 52% for level

1 duration values to 41% for level 5 duration values. Results from an ANOVA showed that the change was significant for all three parameters: F0 ($F(4,25) = 5.467$ $p < 0.001$), intensity ($F(4,25) = 15.86$ $p < 0.001$) and duration ($F(4,25) = 4.2271$ $p < 0.001$). This suggests that the NG subjects use all three acoustic properties for their lexical-stress judgments.

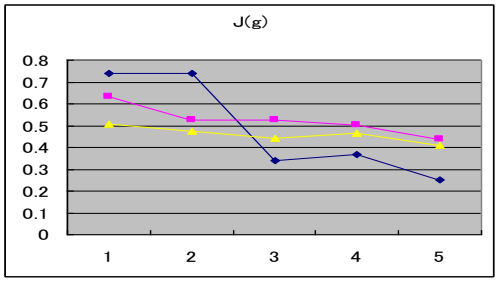
Figure 2 shows that pitch-dependent judgments change greatly in the NJ-group from 75% in stimulus 1 to 30% in stimulus 5, but intensity-dependent judgments vary only between 53% and 47% and duration-dependent judgments change only from 54% to 50%, which is negligible compared to pitch. The ANOVA results show a highly significant effect for F0 ($F(4,95) = 15.51$ $p < 0.001$) but not for intensity ($F(4,95) = 1.589$ $p = 0.15$) nor duration ($F(4,95) = 0.7$ $p = 0.09$). The result supports Beckman's [1] hypothesis that Japanese rely on F0 in lexical accent perception of disyllable words. We suggest that NJ-subjects don't perceive lexical stress in German by using all the parameters, but by using F0 as the sole criterion for stress perception. I.e., the prosodic perception strategy from L1 is transferred to L2. One year of learning German in Japan doesn't contribute to their development of German-type perception. Such a conclusion is supported in the literature, e.g. [12] and [13] for Chinese. Ueyama [12] concludes that Japanese beginners of English don't use duration for English production because they don't use durational differences on the level of Japanese lexical accent. Japanese are mostly sensitive to F0, but insensitive to intensity and duration.

Figure 2: Results for 3 parameters in the NJ group.



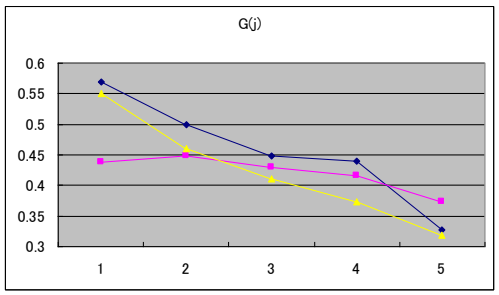
The value of intensity- and duration-dependent judgments in the NJ group are almost the same and remain almost constant from level 1 to level 5, But for the J(g) group, the advanced learners of German who have lived in Germany for more than one year, intensity-dependent judgments change from 63% to 43% and pitch-dependent judgments change from 73% to 25%. In Fig. 3 we see that L1 pitch-dependent behaviour is transferred to L2. The changed sensitivity to intensity must be the result of learning. (ANOVA: F0: ($F(4,45) = 12.448$ $p < 0.001$), I: ($F(4,45) = 3.509$ $p < 0.05$), but D: ($F(4,45)=0.4104$ $p=0.08$)).

Figure 3: Results for 3 parameters in the J(g) group.



The G(j) subjects show the same effects as NG for F0 and duration but it is apparent that the effect of intensity changes, i.e., the effect of intensity disappears. (ANOVA: F0: ($F(4,25) = 3.743$ $p < 0.05$), D: ($F(4,25) = 2.96$ $p < 0.05$), but I: ($F(4.25) = 0.3505$ $p = 0.08$).

Figure 4: Results for the three parameters pitch, intensity and duration of G(j).

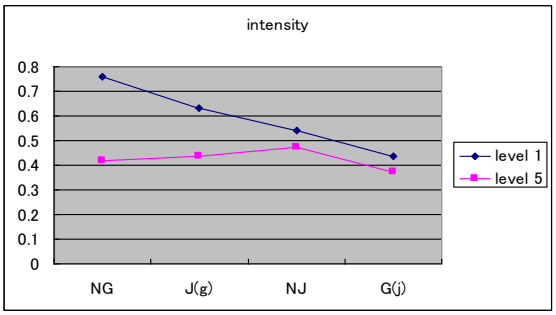


The J(g) and G(j) groups have both adapted their perceptual processing in the direction of the target language (J(g) to German and G(j) to Japanese). They can be said to be using an interlanguage. They have both changed their use of intensity for lexical stress identification. J(g) subjects make greater use of intensity while G(j) subjects make less use of intensity. Interlanguage learners create an interlanguage, using some elements of their native language [9] and other elements from L2. Here both learner groups show the results of their L2 learning in their use of intensity. Fig. 5 shows influence of the intensity parameters for the four groups to make the pattern clearer.

Both NJ and J(g) subjects are sensitive to pitch as a distinctive parameter, but NG and G(j) subjects show a different behavioural pattern (cp. Fig.

2 & Fig. 4). We argue that intensity influences perception in J(g), varying as a function of the period of time the subjects have lived in Germany (76%-41% for NG and 64%-43% for J(g)), while G(j), who are living in Japan, have grown more insensitive to intensity (53%-47% for NJ and 43%-37% for G(j)).

Figure 5: Differences of intensity of J(g), G(j), NJ and NG.



As for duration, the four groups keep it in their native languages. This is because F0 and duration are phonologically processed in Japanese and German as several references point out. Japanese has length contrast and lexical pitch accent, and German length contrast. There is an indication of making use of acoustic cues as prominence in Germany [10]. It suggests intensity influences independently as a phonetic component in both German and Japanese. As PAM (＝Perceptual Assimilation Model) by Best & Tyler [2] point out, when Assimilation of L1-L2 doesn't occur, one or two new L2 phonological categories may be relatively easy to learn perceptually. Also in SLM (＝Speech Learning Model) Flege [7] makes a similar indication to a new phon. This means, F0 and duration have similar functions in L1, they are assimilated into L2. Intensity which is phonetically accessed could be a matter of learning.

3. CONCLUSION

In German, which has lexical stress syllables, and in Japanese, which has pitch accent mora whose prosodic patterns of accent are different, three acoustic correlates of prominence, F0, intensity and duration play a role in the discrimination of lexical stress. NG's perception is based on all the parameters, but NJ, who have been learning German for about one year, only take F0 into account and do not use intensity or duration. This corresponds to the ordinary perceptual patterns of Japanese. J(g), who have had experience living in Germany, show a similar pattern of F0-dependent perception to NJ but their perceptual pattern with regard to intensity has an L2 bias though it differs in degree from NG. G(j), however, who live in Japan, react less to intensity, more like NJ than to NG. The result indicates that F0 and duration are retained because accent is phonemically perceived, but sensitivity to intensity, perceived as a phonetic component, can be developed through teaching or maybe classroom practice.

4. ACKNOWLEDGEMENTS

This study was supported by the Japanese Ministry of Education, Culture, Sports & Technology, Grant-in-Aid for Scientific Research (C) No. 21520953.

5. REFERENCES

[1] Beckman, M. 1986. *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.

[2] Best, C.T., Tyler, M.D. 2007. Non-native and second-language speech perception: Commonalities and complementarities. In Munro, M., Bohn, O.S. (eds.), *Second Language Speech Learning.* Amsterdam: John Benjamins, 13-34.

[3] Boersma, P., Weenink, D. 1992-2003. PRAAT: a system for doing phonetics by computer. *http://www.praat.org*

[4] Dogil, G. 1995. Phonetic correlates of word stress. *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (Univ. Stuttgart), Phonetik AIMS* 2(2), 1-60.

[5] Dogil, G., Williams, B. 1999. The phonetic manifestation of word stress. In van der Hulst, H. (ed.), *Word Prosodic Systems in the Languages of Europe.* Berlin; Mouton de Gruyter, 273-334.

[6] Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.

[7] Flege, J.E. 1995. Second language speech learning: theory, findings, and problems. In Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Baltimore: York Press, 233-277.

[8] Fujisaki, H., Hirose, K., Sugito, M. 1986. Comparison of acoustic features of word accent in English and Japanese. *Journal of the Acoustic Society of Japan* 7, 57-63.

[9] Gass, S., Selinker, L. 1994. *Second Language Acquisition: An Introductory Course*. Hillsdale, NJ; Lawrence Erlbaum Associates.

[10] Jessen, M., Marasek, K., Schneider, K., Classen, K. 1995. Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Proc. 13rd ICPhS* Stockholm 4, 428-431.

[11] Kohler, K.J. 2008. The Perception of Prominence Patterns. *Phonetica* 65, 257-269.

[12] Ueyama, M. 2000. *Prosodic Transfer: An Acoustic Study of L2 English vs. L2 Japanese*. University of California, Los Angeles.

[13] Wang, Q. 2008. L2 Stress Perception: The reliance on different acoustic cues. *Proc. Speech Prosody* Campinas, Brazil, 635-638.