# COORDINATION OF LEXICAL AND PARALINGUISTIC $F_0$ IN L2 PRODUCTION

Yuki Asano

University of Konstanz
yuki.asano@uni-konstanz.de

## ABSTRACT

The present study examines the coordination of lexical and paralinguistic $F_0$ in second language (henceforth L2) production. We tested how L2 learners manage to handle $F_0$ when $F_0$ in their first language (henceforth L1) carries a paralinguistic function, but a lexical one in their L2 and vice versa. Participants (fifteen L1 Japanese speakers and fifteen L1 German speakers, proficient also in their respective L2s) were asked to repeat the same words after an alleged communication failure.

Our results demonstrate negative transfers from the learners' L1 to their L2 in both Japanese and German. The findings are particularly telling, since the words analyzed are very frequent ones, and of which learners should have sufficient L2 input, suggesting that a rich amount of input is not sufficient to the formation of an appropriate $F_0$ in the L2.

**Keywords:** L2 production, lexical and paralinguistic $F_0$, German, Japanese

## 1. INTRODUCTION

$F_0$ is used in different ways across languages. For instance, in Japanese a pitch accent is used lexically and is an inherent property of a word [30]. In German, instead, a pitch accent is not used for lexical distinction, but local $F_0$ movement (=intonation) conveys paralinguistic information [7, 18, 31]. In this study, paralinguistics refers to a speaker's emotion and intention excluding non-linguistic features such as those that cannot be used intentionally [25].

Our study aims at testing how L2 learners, whose L1 and L2 exhibit differences in the use of $F_0$ like those found between Japanese and German, manage to coordinate lexical and paralinguistic $F_0$ to produce L2 utterances in an appropriate way. To this end, we conducted a semi-spontaneous production experiment in which participants produced the same words three times in a row until they succeeded in getting the attention of an imaginary waiter in a crowded and noisy bar. In such a situation, an attitudinal change due to presumably increasing frustration is expected, because the utterance was not heard by the interlocutor. We analyzed whether the deviation of L2 speakers' productions from the L1 speakers' ones is further aggravated when paralinguistic prosody comes into play.

Prosodic adaptations for a paralinguistic purpose have been studied mostly in terms of hyperarticulation [23, 19, 27]. The term hyperarticulation covers a wide range of articulatory adaptations under different intentional, interpersonal or environmental factors. Despite different experimental situations that focused on different languages, previous studies on hyperarticulated speech show similar results such as a slower speaking rate [21, 23, 27]. Regarding Japanese, the coordination of lexical and paralinguistic $F_0$ has been studied in the field of infant-directed speech [15, 16, 22] or corrective focus [20]. Previous studies agree that Japanese L1 speakers do not phonologically change a pitch accent type in hyperarticulated forms. The lexical restriction of the Japanese pitch accent seems to outweigh the modification of local $F_0$ movement. As for German, lexical and paralinguistic uses of pitch do not compete with each other, because pitch is not used lexically, but primarily convey emotion or attitudinal states of a speaker [3, 12, 18] or signal syntactic information such as topic vs. focus, sentence mode (question vs. statement) [6, 10]. For instance, rising pitch accents and boundary tones are known to signal politeness or friendliness [4].

Based on these findings and the theoretical language differences between Japanese and German, we stated the following hypotheses: Japanese L1 speakers will not phonologically change a Japanese pitch accent in the repeated utterances, while German L1 speakers will phonologically vary $F_0$ contours, because this cue conveys the speaker's attitude and emotion in German. Regarding the performance in L2, the way in which $F_0$ is used in the learners' L1 is expected to interfere with the production of their L2's $F_0$. German L2 speakers will fail to be faithful to the lexically fixed defined Japanese pitch accent forms and to vary them. Japanese L2 speakers will not change pitch accent forms due to the lexical restriction in their L1. Besides the analysis of

$F_0$, we additionally analyzed the change of speaking rate in the repeated utterances by measuring total durations of the one-word utterances. As for total durations, we hypothesized that both L1 and L2 speakers' groups will show longer total durations in the repetitions.

## 2. EXPERIMENT

### 2.1. Methods

#### *2.1.1. Participants*

Fifteen speakers of Tokyo-Japanese who were learners of German (8 females, aged between 19 and 36, mean age = 25.1) and 15 speakers of Standard German who were learners of Japanese (6 females, aged between 22 and 35, mean age = 28.9) participated in the experiment in their L1 and L2.

#### *2.1.2. Materials*

The target words used in the study were very frequent Japanese and German words, *sumimasen* ([su.mi.ma.se.ɴ]) and *Entschuldigung* ([ɛnt.ˈʃʊl.di.ɡʊŋ]). Both words mean *excuse me* and can be used in the same context, for instance in calling someone's attention. *Sumimasen* contains a lexically specified pitch fall associated with the penultimate mora in the word, [se].

#### *2.1.3. Design and Procedure*

Following the procedure outlined in Prieto and Roseano [24], materials were presented with descriptions of short scenes. The task was to produce the target words in the given context written in the participants' L1 with a picture of a typical situation:

**Slide 1** You are in a crowded noisy bar. Please call a waiter by saying <*Entschuldigung/ Sumimasen*>.

**Slide 2** He did not hear you. You are a little bit frustrated. Please try it again by saying <*Entschuldigung/ Sumimasen*>.

**Slide 3** He still did not notice it. You are very frustrated by now. Please try it the last time by saying <*Entschuldigung/ Sumimasen*>.

**Slide 4** He finally heard you and is coming to you. Congratulations!

The experiment was designed with Microsoft PowerPoint 2008 and was presented on a Macintosh G3 laptop. In this way, we recorded three attempts for each of the two words in the two languages of each speaker (90 utterances in total). Moreover, 16 filler contexts were provided, in which participants were asked to produce 8 Japanese and 8 German words or short sentences in various contexts. In the first half of the experiment, participants produced utterances in their L1 and in the second half in their L2. Participants were tested individually in a quiet room. Japanese participants were tested in Japan and German participants in Germany. Their responses were digitally recorded onto a computer (44.1kHz, 16Bit) using an unidirectional short-range microphone. In total 180 utterances were used for the analysis (90 each for *sumimasen* and for *Entschuldigung*).

#### *2.1.4. $F_0$ extraction and annotation*

$F_0$ contours were computed using the $F_0$ tracking algorithm in the Praat toolkit [5]. $F_0$ ranges were manually inspected and corrected if there were obvious errors such as octave jumps. Then, segmental boundaries were marked using Praat applying standard segmentation criteria [28].

A manual annotation was carried out by two coders blind to language group and number of attempt. One coder was a Japanese native speaker learning German as a L2 (highly) proficient in German. The other coder was a German native speaker learning Japanese as L2. They annotated the types of accents and boundary tones. Since the variation of pitch accent categories is restricted to H*+L in the Japanese J-ToBI system [29], but German learners of Japanese were expected to produce other accent types influenced by their L1, we used the pitch accent categories in the German G-ToBI system [14] for both the Japanese and German data. The available accent types were six basic pitch accents (H*, L*, L*+H, L+H*, H+L*, H+!H*) [4]. Additionally, the H tones can be downstepped, which increases the inventory from 6 to 11 accents (ibid.). The available boundary tones were L-%, L-H%, H-% and H-^H%.

### 2.2. Results

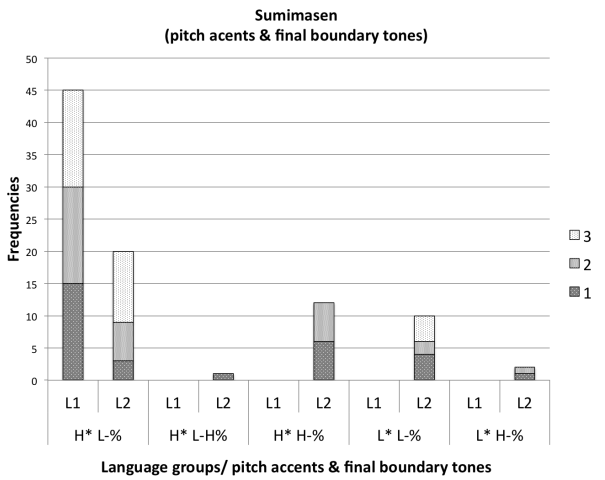
In the following section we will primarily focus on the change between L* and H* pitch accents and L% and H% boundary tones, and regard other variations (e.g. downsteps or upsteps) as secondary modifications. The results from the Japanese native coder will be used as the main source. Similarly, the German native coder will be treated as the principal source for the German utterances.

### 2.2.1. *Results of $F_0$*

***Sumimasen*:** The interrater reliability score for *sumimasen* produced by L1 speakers (Japanese) had a Kappa of 1.00 (SE = 0) and for those produced by L2 speakers (German learners of Japanese) a Kappa of 0.59 (SE = 0.08). The former Kappa value signals an extraordinarily high level of agreement, while the latter shows a moderate level [17].

All 45 L1 speakers' utterances were coded as H* L-%. L1 speakers changed neither pitch accents nor boundary tones across repetitions. In the L2 speakers' 45 data, 20 utterances were coded as H* L-% as the L1 speakers did, while other utterances showed flat contours (=H* H-%) or rising final boundary tones. Contrary to the Japanese L1 speakers' data, German L2 speakers showed variations both in pitch accents and final boundary tones. Noticeably, H% final boundary tones were found only in the first or second attempts. Looking at pitch accent and final boundary tone changes produced by the same speaker across attempts, twelve L2 speakers changed either the pitch accent (7 of them) or the boundary tone (12 of them).

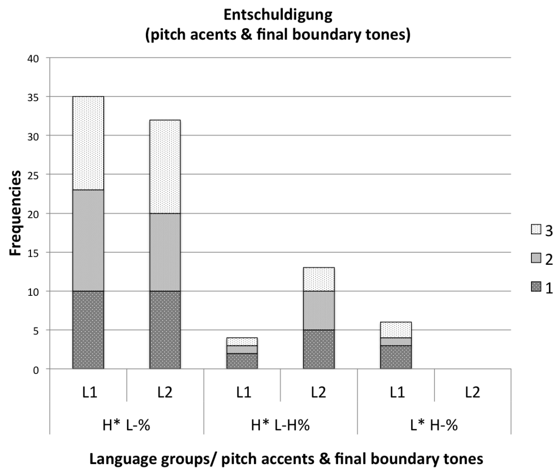
**Figure 1:** Frequencies of pitch accents & final boundary tones for the word *sumimasen* in each attempt produced by L1 and L2 speakers. The legend (1,2,3) refers to the number of attempts.




***Entschuldigung*:** The inter-rater reliability score for *Entschuldigung* produced by L1 speakers (Germans) showed a Kappa of 0.82 (SE = 0.06), and for those produced by L2 speakers (Japanese learners of German) a Kappa of 0.34 (SE = 0.07).

Thirty-five L1 speakers' utterances were coded as H* L-%. The other contours had rising final boundary tones or rising pitch accents. Rising contours occurred more frequently in the first attempt, followed by the second and finally by the third attempt. Regarding the L2 speakers' data, 32 contours were coded as H* L-%, as the same form produced by the L1 speakers. The other utterances showed only a falling pitch accent (H*) followed by a rising final boundary tone (L-H%). No rising pitch accent (L*) has been found. The occurrence of rising pitch accents or rising final boundary tones tends to be more frequent in the first and second attempts overall. Considering the changes of the pitch accent and the final boundary produced by the same speaker across attempts, eleven L1 speakers and seven L2 speakers changed either the pitch accent or the boundary tone.

**Figure 2:** Frequencies of pitch accents & final boundary tones for the word *Entschuldigung* in each attempt produced by L1 and by L2 speakers. The legend (1,2,3) refers to the number of attempts.




### 2.2.2. *Results of total durations*

***Sumimasen*:** Total durations were analyzed using a linear mixed effects regression model with *total durations* as dependent measure and *language group* (L1 vs. L2 speakers) and *number of attempt* (1st., 2nd. and 3rd. attempt) as fixed factors and *participants* as a random factor including random slopes for the fixed factors [2, 8]. Results showed a main effect of *number of attempt* (utterances became longer in repetitions, $p < 0.01$) and of *language group* (L2 speakers generally produced longer utterances than L1 speakers, $p < 0.001$).

***Entschuldigung*:** In the same statistical way, we found a main effect of *number of attempt* (utterances became longer in repetitions, $p < 0.001$) and of *language group* (L2 speakers generally produced longer utterances than L1 speakers, $p < 0.001$).

## 3. DISCUSSION

We investigated the coordination of lexical and paralinguistic $F_0$ in Japanese and German L1 and L2 production. The changes in $F_0$ contours were analyzed by manually annotating pitch accents and boundary tones.

We predicted that German L1 speakers would modify overall $F_0$ contours in repetitions whereas Japanese L1 speakers would not, and that this would also hold true when speaking in their L2. Our results showed that Japanese L1 speakers never changed pitch accents nor boundary tones across the all attempts. They remained, irrespective of the repetitions, faithful to the falling lexical pitch accent associated with the word. German L1 speakers varied both pitch accents and boundary tones in the repetitions. As for L2 productions, German L2 speakers also varied the type of pitch accents and boundary tones in the repetitions. They ignored the lexical restriction in Japanese and modified $F_0$ contours to convey paralinguistic information, as they were used to do it in their L1. Note that they produced deviant forms already in the first attempt, suggesting that it is difficult to produce the lexical pitch accent even without adding the paralinguistic information. Finally, Japanese L2 speakers varied only boundary tones, but not pitch accents. Their faithfulness to the lexical pitch accents in their L1 also influenced their L2 productions. Taking the German L1, German L2 and Japanese L2 groups into account, final boundary tones were more frequently changed than pitch accents. This finding indicates that boundary tones are more likely to carry paralinguistic information than pitch accents, even though pitch accents are more often described to convey information structure in German [4, 13].

The way German L1 speakers varied $F_0$ contours in the repetitions is in line with the current knowledge about the paralinguistic use of German $F_0$ [18, 31]. We found rising boundary tones and pitch accents mostly in the first and second attempt signalling politeness or friendliness. The same result was found in their L2 performance. In the repetitions, participants tried to pursue an response in an unsuccessful communication with falling $F_0$ contours which convey frustration.

The only unexpected result was that Japanese L2 speakers varied boundary tones in the repetitions. We therefore further analyzed the German L1 and Japanese L2 speakers' boundary tones and found that Japanese L2 speakers' H% boundary tones were always preceded by an $H^\star$ L-. This phonological form is an existing one in Japanese to signal interrogatives [11], suggesting that they applied a strategy to produce Japanese interrogatives to call a waiter.

Japanese L1 speakers did not change $F_0$ in the repetitions, but they lengthened their utterances. And this prosodic adaptation was also true for the all participants' groups. Thus, the finding in the previous studies on hyperarticulation was confirmed. The analysis of total durations additionally revealed that the utterances produced by L2 speakers were generally longer than those produced by L1 speakers. Slower speech rate in L2 than in L1 has been reported in other studies [1, 9].

Finally, we reported the inter-rater reliability scores. The following order of the scores was found: the score for the Japanese L1 speakers was greater than those for the German L1 speakers followed by those for the German L2 speakers, and finally by those for the Japanese L2 speakers. The scores for the L1 speakers' data were generally higher than those for the L2 speakers' ones. The comparison between *sumimasen* and *Entschuldigung* shows higher values for the former utterance than for the latter. The greater the variation of pitch accents and boundary tones was, the poorer were the scores of the agreement, showing a general practical problem in annotating variegated data. This finding indicates that the annotation of L2 learners' data was more difficult than that of L2 speakers'. It is because their pitch accents and boundary tones were difficult to categorize into the existing categories in a ToBI system. The L2 data are characterized not only the mixture of L1 and L2 prosody, but they also show the dynamic character of an interlanguage [26].

Taken together, our results clearly confirm the language-specific ways to modify $F_0$ for signalling paralinguistic meaning and the negative transfer from one's L1 to L2. Our findings are especially noteworthy as they were found in highly frequent words in Japanese and German that learners should have encountered very often, suggesting that a rich amount of the input of the target language did not contribute to the formation of an appropriate L2 prosody. As the next step, we are currently conducting perception and production experiments to determine the source of these deviant L2 production forms.

## 4. REFERENCES


[1] Anderson-Hsieh, J., Koehler, K. 1988. The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning* 38(4), 561–613.

[2] Barr, D. J., Levy, R., Scheepers, C., Tily, H. J. 4 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3), 255–278.

[3] Baumann, S., Grice, M. 2006. The intonation of accessibility. [invited contribution to journal of pragmatics]. *Journal of Pragmatics* 38, 1636–1657.

[4] Baumann, S., Grice, M., Benzmüller, R. 2001. GToBI – a phonological system for the transcription of German intonation. In: Puppel, S., Demenko, G., (eds), *Prosody 2000: Speech recognition and synthesis* 21–28. Adam Mickiewicz University.

[5] Boersma, P., Weenink, D. 2011. Praat: doing phonetics by computer [computer program] version 5.2.20.

[6] Braun, B. 2006. Phonetics and phonology of thematic contrast in german. *Language and cognitive processes* 26, 224–235.

[7] Chen, A. 2005. *Universal and language-specific perception of paralinguistic intonational meaning*. LOT.

[8] Cunnings, I. 2012. An overview of mixed-effects statistical models for second language researchers. *Second Language Research* 28(3), 369–382.

[9] Derwing, T. M., Munro, M. J. 1997. Accent, intelligibility, and comprehensibility: Evidence from four l1s. *Studies in Second Language Acquisition* 20, 1–16.

[10] Féry, C. 1993. *German Intonational Patterns*. Tübingen: Niemeyer.

[11] Fujisaki, H., Hirose, K. 1993. Analysis and perception of intonation expressing paralinguistic information in spoken japanese. *ISCA Workshop on Prosody* 254–257.

[12] Gibbon, D. 1998. Intonation in German. In: Hirst, D., Cristo, A., (eds), *Intonation systems: a survey of twenty languages*. Cambridge University Press chapter 4, 78–95.

[13] Grice, M., Bauman, S. 2007. An introduction to intonation - functions and models. In: Trouvain, J., Gut, U., (eds), *Non-Native Prosody: Phonetic Description and Teaching Practice*. Mouton De Gruyter chapter 1, 25–52.

[14] Grice, M., Baumann, S., Benzmüller, R. 2005. German intonation in autosegmental-metrical phonology. In: Sun-Ah, J., (ed), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press 55–83.

[15] Ishihara, S. 2011. Japanese focus prosody revisited: Freeing focus from prosodic phrasing. *Lingua* 121(13), 1870–1889.

[16] Kitamura, C., Burnham, D. 2003. Pitch and communicative intent in mother's speech: Adjustment for age and sex in the first year. *Infancy* 41(1), 85–110.

[17] Landis, J. R., Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.

[18] Liscombe, J. 2007. *Prosody and Speaker State: Paralinguistics, Pragmatics, and Proficiency*. PhD thesis Columbia University.

[19] Lombard, É. 1911. Le signe de l'élévation de la voix. *Annales des maladies de l'oreille et du larynx* 37, 101–119.

[20] Maekawa, K. 2004. Production and perception 'paralinguistic' information. volume 03. Proceeding of Specch Prosody.

[21] Moon, S.-J., Lindblom, B. 1994. Interaction between duration, context, and speaking style in english stressed vowels. *Journal of Acoustical Society of America* 96, 40–55.

[22] Nagahara, H. 1994. *Phonological phrasing in Japanese*. PhD thesis University of California Los Angeles.

[23] Oviatt, S., Levow, G. A., Moreton, E., MacEachern, M. 1996. Modeling global and focal hyperarticulation during human-computer error resolution. *J Acoust Soc Am* 104(5), 3080–98.

[24] Prieto, P., Roseano, P., (eds) 2010. *Transcription of Intonation of the Spanish Language*. Munich: Lincom Europa.

[25] Roach, P., Stibbard, R., Osborne, J., Arnfield, S., Setter, J. 1998. Transcription of prosodic and paralinguistic features of emotional speech. *Journal of the international Phonetic Association* 28, 83–94.

[26] Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics* 10, 209–241.

[27] Stent, A. J., Huffman, M. K., Brennan, S. E. 2008. Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication* 50(3), 163–178.

[28] Turk, O., Schöder, M., Bozkurt, B., Arslan, L. 2005. Voice quality interpolation for emotional text-to-speech synthesis. *Proceedings of Interspeech* Lisbon. 797–800.

[29] Venditti, J. 1995. Japanese ToBI labelling guidelines.

[30] Venditti, J. 2000. *Discourse Structure and Attentional Salience Effects on Japanese Intonation*. PhD thesis The Ohio State University.

[31] Wichmann, A. 2000. The attitudinal effects of prosody, and how they relate to emotion. *ISCA Workshop on Speech and Emotion*. Belfast, Northern Ireland.