The Perception-Production Link in Intonation:
Evidence from German Learners of English

Karin Puga[1], Robert Fuchs[2], Toby Hudson[3], Jane Setter[3], Peggy Mok[4]

[1]Justus Liebig University Giessen, Germany,
[2]University of Hamburg, Germany
[3]University of Reading, United Kingdom,
[4]The Chinese University of Hong Kong, Hong Kong SAR

Karin.Puga@anglistik.uni-giessen.de, robert.fuchs@uni-hamburg.de, toh22@cam.ac.uk,
j.e.setter@reading.ac.uk, peggymok@cuhk.edu.hk

Abstract

Investigations of the link between the perception and production of prosody by language learners can inform theories of prosody perception and production, especially with regard to Second Language Acquisition (SLA), and for the implementation of prosody in Foreign Language Teaching (FLT). The perception and production of prosody in L2 speech are often analyzed separately, but the link between the two is rarely the focus of investigation [e.g. 1, 2].

In a previous study [3], we analyzed the perception of prosody in read speech by German learners of English (n=20), who performed similarly to the British English (BrE) control group (n=25) for some sentence types (e.g. statements, yes/no-questions) and worse for others (e.g. open and closed tag questions, sarcasm). The present study extends this analysis by comparing the same learners' perception and production of prosody in read speech with the same sentence types.

Overall, the learners (n=20) performed better in production and were more similar to the native speakers' (n=10) performance than in the perception task. However, the learners significantly differed from the native controls in production, i.e. closed tag questions and checking questions. Interestingly, the learners also performed significantly better in yes/no and statement questions than the native speakers.

Index Terms: L2 intonation, L2 acquisition, English, German, perception vs. production

1. Introduction

Prosody is a particularly challenging area in L2 speech learning, even at the advanced level, as previous studies on L2 prosody have shown [4-8]. The reasons for this include that prosody is highly dependent on context and fulfills various functions (e.g. attitudinal, discoursal), some of which the learners might not be aware of. For learners undergoing formal instruction, the problem is exacerbated by a general neglect of prosody in FLT [see e.g. 9, 10].

Previous studies of L2 production show that speech produced by learners of English differs significantly from the speech produced by native speakers [1, 8]. Studies of L2 perception, however, reported only a few problematic areas for L2 speakers, such as tag questions and sarcasm [3, 13]. This might suggest that, even though learners of English are able to recognize and understand intonational patterns, their production may lag behind their abilities in speech perception.

Current L2 speech learning theories stress the importance of an interaction between perception and production in the successful acquisition of L2 phonology [14, 17, 18]. More specifically, theories such as the Speech Learning Model [14], while primarily concerned with segmental rather than supra-segmental features, posit that correct perception is a precondition to correct production [e.g. 15]. However, other research has suggested that accurate production may precede accurate perception of nonnative prosody [16, 17], and that the perception-production link may be more complex depending on a variety of factors (e.g. phonetic, phonological, social, etc.). Therefore, the present study investigates the following research questions:

(1) What intonational patterns do L2 learners of English produce depending on sentence type?

(2) In which respects do the learners differ from native speakers in their production in similar contexts?

(3) Is there an implicational relationship between the perception and production of prosody by L2 learners such that correct perception implies correct production but not the inverse?

By answering these research questions, the present study aims to explore the link between prosody perception and production. We analyze the ability of German learners of English to choose appropriate intonational patterns for various sentence types (e.g. questions, statements, sarcasm, etc.) that are embedded in a narrative context and compare this with the results of our previous perception study [3].

2. Studies on L2 Perception & Production

The small genetic and typological distance between German and English (both are intonation languages with a stress-timed rhythm and both use pitch in a systematic way to mark different syntactic structures [20-23]) suggests only limited differences in tone inventory and the meaning of tones [20, 24-26]. Five mono- and bi-tonal nuclear tones (rise, fall, fall-rise, rise-fall, and level) have been described in the British tradition across various studies [13, 27] for English as well as German. Even though studies have shown that German and English seem to share a common tone inventory of phonological representations, the two languages differ in the way these representations are realized phonetically, i.e. the alignment of the peak might be realized on different vowels or for a longer duration, etc. [24, 28].

Production studies indicate that in BrE falling tones are used 50% of the time and rises and fall-rises about 40% of the time [27]. In German learner English, level tones are the most common type of nuclei (65.4% of all cases) [8]. Simple pitch movements are distributed equally with falling (17.6%) and rising nuclei (15.5%), while complex tones are extremely rare in nonnative speech (<1.5%) [8]. Also, an overuse of rises and replacement of falls with rises and vice versa were reported for German learners of English [8].

Another interesting aspect is the acquisition of nuclear tones. A longitudinal study of nuclear tones produced by Austrian school children learning English in a reading passage showed that simple tones are acquired before complex tones, and falls are produced before rises [29]. Native speakers of English of the same age produced the same proportions but the falls were phonetically different. Additionally, the German learners produced only 17.5% of rises in English and 45.9% of rises in the same context in their L1. Thus, intonation training seems to have the greatest influence on the acquisition of intonational patterns, as other studies have shown that German English (GE) speakers significantly increased the overall percentage of nuclei with pitch movements and complex nuclei similar to native speaker values in the reading passage after a six-month training course [8].

While a sizeable number of studies have investigated the production of prosody, very few studies have focused on perception. The two studies most relevant to the present analysis were conducted by Mok et al. (Hong Kong English (HKE) vs. BrE) [13] and Puga et al. (GE vs. BrE) [3]. Both studies tested the learners' knowledge of nuclear tones in an intonation perception task and compared them to a native control group. The GE learners chose the correct intonation contour slightly more than half of the time (54.1%), which was more often than the HKE learners at 43%, but less often than the BrE speakers (72.6%). The results showed that the two learner groups were quite similar to each other. Both learner groups encountered problems with tag questions and sarcasm. These results were partly explained by a possible L1-influence.

In summary, for L2 speech learning in general, a tight link between production and perception has been postulated by many authors [14, 17, 30-33]. While some studies have shown that adult-like production of prosody precedes comprehension [16, 17], the more common pattern appears to be that perception precedes production [34, 35]. Most of the difficulties L2 learners experience seem to be motivated perceptually [e.g. 15] and often related to L1-transfer [1, 8]. Morever, some studies suggest that L1-transfer is more evident in production than in perception, further supporting the claim that perception precedes production [1].

3. Method

3.1. Participants

20 German learners of English (5 male; age 20-28, mean 24), all L1 speakers of German with no speech or auditory impairments, participated in the experiment. All of the participants were advanced learners of English enrolled in a Bachelor's or Master's degree in English literature and linguistics. 6 of the 20 participants had stayed in an English-speaking country for a period ranging from one month to one year (see [3] for further information).

The BrE native speaker control group consisted of 10 participants (3 male) with normal speech and hearing. They were all university students aged 19 to 34 (mean 22.5 years; SD: ±1.7).

3.2. Data

Participants were asked to read aloud a short story with a length of 523 words/28 sentences that included several sentence types (e.g. statements, questions, sarcasm). The expected nuclear tones (see Table 1) for each sentence type in the short story are based on the fourth author's initial reading and standard descriptions of BrE [13, 27]. In the case of sarcasm, we distinguished between two different types of sarcasm contours. Sarcasm 1 (appears twice in the experiment) could be produced with a rise-fall or a fall. Sarcasm 2 (appears once) could be produced with a rise-fall only.

Table 1: *Number of items by sentence type and expected nuclear tone.*

Sentence Type	Nuclear tone	Number of items
Statement	fall	3
Continuation	level/rise	3
Statement question	rise	3
Echo question	rise/fall-rise	3
Yes/no question	rise	3
Wh-question	fall	3
Closed tag	fall	3
Open tag	rise	2
Checking tag	rise	1
Sarcasm	rise-fall/fall	3
Checking	fall-rise/rise	1

3.3. Procedure

The participants were first allowed to familiarize themselves with the short story and were then recorded reading it aloud. They were instructed to read the story in its entirety and not to stop for self-corrections. After the production task, the participants completed the perception task on the basis of the same short story ([3]). The recordings were annotated in Praat for nucleus placement and nuclear tone contours.

3.4. Statistics

The data were analyzed with a logistic mixed-effects regression model with RESPONSE (expected/other) as dependent variable. The model used in section 4.1 included an interaction between L1 (English/German) and SENTENCE_TYPE as independent variables as well as PARTICIPANT and ITEM as random effects (computed with packages NLME and LSMEANS in R [36-38]). The models in section 4.4 included SENTENCE_TYPE as independent variable and PARTICIPANT and ITEM as random effects. Significance testing was carried out with post-hoc Tukey tests.

4. Results

4.1 Differences and similarities in the production data

The learners in this study reached only a slightly lower accuracy rate (mean 67.1%) than the native control group (mean 70.7%; see Table 2). Nevertheless, an analysis by sentence type reveals substantial differences between the two groups. Specifically, a mixed effects regression model on accuracy of tone choice indicates that the learners performed significantly worse than

the native group in closed tag and checking questions, and significantly better in yes/no and statement questions. For six out of thirteen sentence types, the learners reached accuracy rates similar to the control group, including checking tag, statements, open tag questions, wh-questions, sarcasm 2, and continuation. The accuracy rates in Table 2 are based on the fourth authors' recording and the expected contours for certain sentence types determined by a literature review [13, 27]. If the tone choice made by the native control group is used as a baseline instead, the accuracy rate of all individual learners drops by about 20%, ranging from 30% to 60% (mean 47%).

Table 2: *Distribution of choices for each sentence type* (rounded percentages; highlighted cells are the expected answers; G=German learners; E=English native speakers; Q=Question). Statistical significance of differences in accuracy (correct vs. incorrect tone) between G and E are indicated under *Type*.

Type (Tukey test)	Group	Fall	Rise	Fall-rise	Rise-fall	Level
Statement	G	80	8	6	7	0
	E	93	0	3	0	3
Continuation	G	38	8	8	0	43
	E	40	3	27	0	30
Statement Q (p<0.001)	G	10	77.5	7.5	5	0
	E	5	15	80	0	0
Statement question/echo	G	5	90	5	0	0
	E	10	10	80	0	0
Echo question	G	3	87	10	0	0
	E	3	37	57	0	3
YN-question (p<0.0001)	G	5	95	0	0	0
	E	13	23	63	0	0
Wh-question	G	77	18	1.6	1.6	1.6
	E	93	0	3	3	0
Closed tag (p<0.0001)	G	25	68	2	2	3
	E	90	3	3	0	3
Open tag	G	10	80	7.5	2.5	0
	E	5	95	0	0	0
Checking tag	G	0	100	0	0	0
	E	0	100	0	0	0
Sarcasm (1)	G	12.5	37.5	7.5	42.5	0
	E	60	0	5	35	0
Sarcasm (2)	G	75	15	0	5	5
	E	60	0	20	20	0
Checking (p<0.01)	G	70	5	20	5	0
	E	10	0	90	0	0

The two groups show similar performance in several contexts, including in checking tag sentences, where both groups only produced rises, which was also the expected tone. Both groups (GE and E) produced similar tones with almost equal accuracy rates for simple statements (fall= BrE=93%; GE=80%). For open tag sentences, both groups produced mostly rises (GE=80%; BrE=95). Both groups were mostly accurate in choosing a falling tone for wh-questions (GE=77%; BrE=93). For sarcasm 2, where only a rise-fall tone was deemed possible, the native speakers outperformed the learners (E=20%;

GE=5%). However, both groups opted mostly for falls in this condition (GE=75%; BrE=60%), which was not one of the expected contours. Continuation sentences were also produced similarly among the two groups with expected (level) and not expected contours (falls).

Apart from these similarities, a number of substantial differences were found, too. Interestingly, the German learners significantly outperformed the native speakers in two conditions, i.e. yes/no-questions (GE=95%; BrE=23%), where (in addition to the canonical rise) the native speakers also produced fall-rises (63%) and falls (13%; 5% of German learners also produced falls). The second condition included statement questions, where simple rises were expected, too (GE=77.5%; BrE=15%): here the native speakers produced fall-rises (80%) and falls (5%). On the other hand, the checking and closed tag sentences seemed to pose the largest problem in production for the nonnative speakers of English. For the checking sentence, 90% of the native speakers produced a fall-rise nuclear tone, but only 20% of the nonnative speakers did so. While both groups reached similar accuracy rates for statement questions/echo (GE=95%; BrE=90%) and echo questions (GE=97%; BrE=94%), the distribution among the two possible tones for both sentence types was different.

Overall, the native speakers seemed to be more consistent as a group in the production of sentence types where falls were expected and where nonnative speakers quite frequently opted for rises instead (e.g. wh-questions, statements, sarcasm 1).

4.2 Differences between individual speakers

Accuracy rates for most of the British participants ranged from 64% to 70%, with only two speakers reaching 78.6% and 85.7%, respectively. Unsurprisingly, the learners were more heterogeneous in their performance than the native speakers, with accuracy rates ranging from 42.9% to 85.7%, and most learners in the 60-70% range.

Looking more closely at the learner biographies of the eight best-performing participants (accuracy above 70%), we find no indication that particular variables might explain their good performance. The eight learners comprise both younger and older participants, were mostly brought up monolingually and some but not all of them spent time in an English-speaking country (USA and New Zealand). The only biographical variable providing some insight is gender, with all eight best-performing learners being female. However, this finding should be interpreted with caution given that only five of the twenty learners are male, and all five reached average scores.

4.3 Alternative nucleus placement

Every learner in the present data set committed at least two nucleus placement errors, and some learners up to six (GE: n=87 nucleus placement errors, BrE: n=25 nucleus placement errors). The five most frequent nucleus placement errors occurred in the two sarcasm sentences, one checking sentence, one statement, and one wh-question. For instance, common errors for the nonnative group appeared in the sarcastic remark *Would you, now? Luke said sarcastically*, where 11 of the 20 learners placed the nucleus on *now* instead of *would you* (and one learner on *both*). Trials with erroneous nuclear placement were not included in the subsequent analysis.

4.4 Comparison of accuracy in production and perception

Overall, we find no evidence of an implicational relationship between correct perception and correct production. We compared the tone choice in the perception and production tasks for each learner across contexts and distinguish four scenarios: (1) correct choice in both tasks, (2) incorrect choice in both tasks, (3) correct choice in perception task only, and (4) correct choice in production task only. If correct production requires correct perception, scenario 4 should be rare or at least rarer than scenario 3.

The analysis reveals that the learners chose the correct tone in production only (scenario 4) almost twice as often as in perception only (scenario 3), as shown in Table 3 ('full dataset'). Given that even the native control group performed badly in some contexts (see section 4.1), we removed the four sentence types with the worst scores for the control group from the analysis (continuation, statement qn., y/n qn., sarcasm). Even in this reduced dataset, the learners still chose the correct tone more often in production only than in perception only.

Table 3: *Relationship between performance in production and perception tasks, learner data only.*

Correct	Full dataset		Reduced dataset	
	N	%	N	%
(1) Both	224	40.0	173	50.9
(2) None	105	18.8	51	15.0
(3) Perception only	79	14.1	50	14.7
(4) Production only	152	27.1	66	19.4

Returning to the full dataset, we further investigated whether accurate choice in scenario (4), production only, and scenario (1), correct perception and production, was more likely to occur for specific sentence types. A mixed effects regression model confirms this and indicates that accurate choice in production only was particularly likely to occur in yes/no-questions (estimated likelihood .58), continuation (.42), checking tag questions (.40) and open tag questions (.33). Correct choice in both perception and production was particularly likely in echo questions (.72), statement questions/echo (.70), statements (.62) and wh-questions (.60).

5. Discussion

Based on an analysis of intonation contour choice embedded in a short story context (research question 1), we find that the 20 L1 German learners of English tested in this study performed better in the production task than in the perception task. While their performance in the perception task showed similarities with the native control group, substantial differences were found in production, especially in particular sentence types (e.g. closed tag questions and checking questions). Interestingly, the learners also performed significantly better than the control group in two contexts (yes/no and statement questions). A plausible explanation might be that the learners tend to adopt standard BrE intonation contour choices, while the native control group might follow a more recent model that is the result of ongoing language change not yet reflected in English Language Teaching. At present, this explanation remains speculative and should be tested in future work.

As for the choice of particular non-target like intonation patterns, just as previous studies have shown [8], the learners produced falls where rises would have been expected, and vice versa. This pattern was found in both perception (statements, w-questions, tag questions) and production (e.g. sarcasm, wh-questions).

Comparing the performance of individual learners in the perception task [3] and production task, we find that more participants reached accuracy rates similar to the native control group in the production task (8/20) than in the perception task (4/20). Performance was strongly dependent on sentence type, with some types likely to induce correct performance in both perception and production, some in perception only, and some in production only.

While there are some differences between learners and native speakers in tone choice in particular contexts (research question 2), the learners perform significantly better in production than in perception. This is relevant for our third research question, which asked whether correct perception in learners implies correct production, as speech learning theories predict [14]. Our results indicate the opposite pattern, with the unexpected pattern, correct production without correct perception, in 27% of all trials, and the expected pattern, correct perception without correct production, in only 14%. Thus, correct perception does not appear to imply correct production, but rather the other way around, as some previous studies have suggested [16, 17]. The results further suggest that the perception and production of prosody are aligned in some, but not all contexts. However, it needs to be kept in mind that the native control group did not remain stable in their performance across tasks either. For instance, while the native speakers only reached an accuracy of 44% for checking sentences in the perception data, 90% accuracy was reached in the production task for this sentence type.

When evaluating other influential factors in the production data, we found no indication that particular variables might explain the good performance of individual learners for some sentence types. Only gender might play a role in both perception as well as production, with female learners outperforming the male learners (and more so in perception than production). As in the perception task, L1 influence might account for the problems the learners faced with closed tag questions [13], since tag questions do not have a direct equivalent in German. Apart from subject variables, speaker-independent and contextual variables of the experiment also provide some insight [39, 40]. These variables could account for the unexpected results for the continuation sentences in both the perception and production tasks, where both groups might not have been aware that the sentences were intended as "continuation sentences".

Finally, the analysis of the production data revealed further differences between native and nonnative speech. These might be addressed in a follow-up study. One area of interest is the specific phonetic realization of nuclear tones, including peak alignment. Given that German and English have been described to have different phonetic realizations of several nuclear tones [24, 28, 29], we expect that learners transfer fine phonetic detail from their L1 to their L2 at least to a certain degree.

6. Acknowledgements

7. References

[1] M. Ortega-Llebaria & L. Colantoni "L2 English Intonation," *Relations between Form-Meaning Associations, Access to Meaning, and L1 Transfer.* Studies in Second Language Acquisition, 2014, 36, 331-353, 2014.

[2] Altmann, H. (2006). The perception and production of second language stress: A cross-linguistic experimental study (Unpublished doctoral dissertation). University of Delaware, Newark.

[3] K. Puga, R. Fuchs, J. Setter, & P. Mok, "Perception of Intonation Patterns by German L2 Speakers of English," Stockholm, Sweden, *Interspeech*, 2017.

[4] T. Bongaerts, C. Van Summeren, and B. Planken & E. Schils, "Age and ultimate attainment in the pronunciation of a foreign language," *Studies in Second Language Acquisition* 19, pp. 447-465, 1997.

[5] W. Grosser, "Aspects of intonational L2 acquisition," *Current issues in European second language acquisition research*, eds. B. Kettemann & W. Wieden. Tübingen: Gunter Narr Verlag, pp. 81-94, 1993.

[6] A. Li and B. Post, "L2 Acquisition of prosodic properties of speech rhythm – Evidence from L1 Mandarin and German Learners of English," *Studies in Second Language Acquisition* 36 (2), pp. 223-255, 2014.

[7] T. Scovel, "A critical review of the critical period research," *Annual Review of Applied Linguistics* 20, pp. 213-223, 2000.

[8] U. Gut, *Non-native speech: a corpus-based analysis of phonological and phonetic properties of L2 English and German.* Frankfurt: Lang, 2009.

[9] U. Hirschfeld and J. Trouvain, "Teaching prosody in German as foreign language," *Non-Native Prosody – phonetic description and teaching practice*, eds. J. Trouvain & U. Gut. Berlin: Mouton de Gruyter, pp. 171-187, 2007.

[10] S.M. Gass and L. Selinker, *Second Language Acquisition: An Introductory Course*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates, 2001.

[11] E. Grabe, B.S. Rosner, J.E. García-Albea, and X. Zhou, "Perception of English Intonation by English, Spanish and Chinese listeners," *Language and Speech*, 46, pp. 375-401, 2003.

[12] J. Setter, V. Stojanovik, and P. Martinez Castilla, "Evaluating the intonation of non-native speakers of English using a computerized test battery," *International Journal of Applied Linguistics*, 20, pp. 368-385, 2010.

[13] P. Mok, Y. Yin, J. Setter, and N.M. Nayan, "Assessing knowledge of English intonation patterns by L2 speakers," Boston, MA, *Speech Prosody*, 2016, pp. 543-547.

[14] J. E. Flege. "Second language speech learning: Theory, findings, and problems." In Speech perception and linguistic experience: Issues in cross-language research, ed. W. Strange, 233–277, 1995. Timonium: York Press.

[15] P. Trofimovich, and W. Baker. Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition* 28: 1–30. 2006.

[16] A. Sheldon & W. Strange, "The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception," Applied Psycholinguistics, 19, 1982, pp. 243-261.

[17] T. O. Lentz & A. Chen, "Unbalanced adult production and perception in prosody" ICPhs, 2015.

[18] C. T. Best, "A direct realist view of cross-language speech perception." In Speech perception and linguistic experience: Issues in cross-language research, ed. W. Strange, 171–232, 1995. Timonium: York Press.

[19] I. Mennen. "Beyond segments: Towards a L2 intonation learning theory." In *Prosody and language in contact*, E. Delais-Roussarie, M. Avanzi and S. Herment, eds. pp. 171-188. Springer Berlin Heidelberg, 2015.

[20] J. Esser, "Contrastive intonation of German and English," *Phonetica* 35, 41-55, 1978.

[21] A. Batliner & W. Oppenrieder, *Zur Intonation von Modus und Fokus im Deutschen*, eds. H. Altmann,. Tübingen: Niemeyer. 87-109, 1989.

[22] Oppenrieder, W. (1988a): "Intonation und Identifikation. Kategorisierungstests zur kontextfreien Identifikation von Satzmodi", *Intonationsforschungen*, ed. H. Altmann. Tübingen: Niemeyer. 153-167.

[23] Oppenrieder, W. (1988b): "Intonatorische Kennzeichnung von Satzmodi", *Intonationsforschungen*, ed. H. Altmann. Tübingen: Niemeyer. 169-205.

[24] E. Grabe, "Comparative Intonational Phonology: English and German," *Proceedings of the ESCA Workshop on Intonation: Theory, Models, and Applications*. Athens, Greece, pp. 157- 160, 1998.

[25] A. Fox, *German Intonation: An Outline.* Oxford. Clarendon Press, 1984.

[26] M. Scuffil, (1982): *Experiments in comparative intonation: a case-study of English and German.* Tübingen: Niemeyer.

[27] A. Cruttenden, *Gimson's Pronunciation of English (8th edition).* London: Routledge, 2014.

[28] I. Mennen, "Bi-directional interference in the intonation of Dutch speakers of Greek", *Journal of Phonetics* 32, 543-563. 2004.

[29] W. Grosser, "On the acquisition of tonal and accentual features of English by Austrian learners," in *Second Language Speech: Structure and Process*, ed. A. James & J. Leather. Berlin: Mouton de Gruyter, 1997, pp. 211-228.

[30] L. Goldstein, Fowler, C. A. 2003. Articulatory phonology: a phonology for public language use. In: Schiller, N. O., Meyer, A. S. (eds), *Phonetics and Phonology in Language Comprehension and Production*. Berlin: Mouton de Gruyter, 159-207.

[31] M. Grosvald, 2009. Inter-speaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation. *J. Phonetics* 37, 173-188.

[32] S. Lin, 2011. Perception and production of prosodically varying inter-gestural timing in American English laterals. Ph.D. dissertation, University of Michigan, Ann Arbor.

[33] P. Speeter Beddor, "The Relation between Language Users' Perception and Production Repertoires," ICPhs, 2015, pp. numbers.

[34] A. Chen, 2010. Is there really an asymmetry in the acquisition of the focus-to-accentuation mapping? *Lingua* 120, 1926–1939.

[35] A. Cutler, Swinney, D. A. 1987. Prosody and the development of comprehension. *Journal of Child Language* 14(1), 145–167.

[36] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2013.

[37] J. Pinheiro Bates D, DebRoy S, Sarkar D and R Core Team (2017). _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-131

[38] R. V. Lenth and M. Hervé. lsmeans: Least-Squares Means. R package eversion 2.17. http://CRAN.R-project.org/package=lsmeans, 2015.

[39] M.J. Munro & T.M. Derwing, The effects of speaking rate on listener evaluations of native and foreign-accented speech. Language Learning 48, pp. 159-182. 1998.

[40] S.V. Levi, S.J. Witers, & D.B. Pisoni, "Speaker-independent factors affecting the perception of foreign accent in a second language. Journal of the Acoustical Society of America 121: 2327-2338.