



The Perception of English Intonation Patterns by German L2 speakers of English

Karin Puga,¹ Robert Fuchs,² Jane Setter,³ Peggy Mok⁴

¹Justus Liebig University Giessen, Germany,

²Hong Kong Baptist University, Hong Kong SAR

³University of Reading, United Kingdom,

⁴Chinese University of Hong Kong, Hong Kong SAR

Karin.Puga@anglistik.uni-giessen.de, rfuchs@hkbu.edu.hk, j.e.setter@reading.ac.uk,
peggymok@cuhk.edu.hk

Abstract

Previous research suggests that intonation is a particularly challenging aspect of L2 speech learning. While most research focuses on speech production, we widen the focus and study the perception of intonation by L2 learners. We investigate whether advanced German learners of English have knowledge of the appropriate English intonation patterns in a narrative context with different sentence types (e.g. statements, questions). The results of a tonal pattern selection task indicate that learners (n=20) performed similar to British English controls (n=25) for some sentence types (e.g. statements, yes/no-questions), but performed significantly worse than the control group in the case of open and closed tag questions and the expression of sarcasm. The results can be explained by the fact that tag questions are the only sentence type investigated that does not exist in the learners' L1, and sarcasm is not represented syntactically. This suggests that L1 influence can partly account for why some intonation patterns are more challenging than others, and that contextualized knowledge of the intonation patterns of the target language rather than knowledge of intonation patterns in isolation is crucial for the successful L2 learning of intonation.

Index Terms: L2 intonation, L2 acquisition, English, German, perception

1. Introduction

Intonational patterns are very difficult to master for second language (L2) learners and even advanced learners of English still deviate from native-like tonal patterns [1-4]. Part of the reason why the L2 acquisition of intonation is challenging might be that it fulfills various functions (e.g. attitudinal, discursal, etc.) and is highly dependent on context.

While a substantial body of research has shown that the L2 production of intonation patterns differs substantially from those of native speakers [see, e.g. 5, 6], there is a lack of evidence on the perception of native (L1) English intonation patterns by L2 learners. However, current L2 speech learning theories stress the importance of an interaction between perception and production in the successful acquisition of L2 phonology [7, 8]. The present study, therefore, seeks to answer the following research questions: (1) What knowledge do L2 learners of English possess of appropriate intonation patterns in different contexts? (2) Can their (lack of) knowledge of particular patterns be explained by influence from their L1?

In order to cater for the need for more research on the L2 perception of intonation, the present study focuses on the ability of German learners of English to choose appropriate intonation patterns for a range of sentence types (e.g. questions and statements) embedded in a narrative context.

2. Studies on L2 Intonation

Formal instruction in English starts in the German education system between the age of 7 and 11, and remains an obligatory subject for almost all students until they leave school. English is taught primarily as a foreign language by nonnative speakers of English (L1 German). Prosody is mostly ignored in German schools, partly because the nonnative teachers might themselves be unsure which intonation patterns are 'correct', and might therefore be hesitant to address intonation in formal instruction [9, 10]. While some learners who spend a semester abroad or have contact with native speakers may acquire native-like intonation patterns, most students fail to do so [11, 12].

German and English are both intonation languages with a stress-timed rhythm [13]. Both have been described to have very similar intonation patterns [13-15]. A few studies have been conducted that deal with German English intonation patterns [16]; these were restricted to production data only, and show that German-speaking learners of English do have problems with producing native-like intonation patterns [16].

The most common tone type for British English (BrE) is the falling tone (50%), followed by rise and fall-rise (40%) [17]. Gut [16] found that, in non-native English, level tones are the most common type of nuclei (65% of all cases). Simple pitch movements are distributed equally with falling (18%) and rising nuclei (16%). Complex intonation patterns are reported to be extremely rare in nonnative speech (<1.5%) [16]. Previous studies on sentence types and tone patterns of German English reported that learners overuse rises [16].

Grosser's [18] longitudinal study of adolescent Austrian German learners of English found that simple tones are acquired before complex tones, and falls are produced before rises. Interestingly, after one year of instruction, the learners realized only 18% of all nuclear tones as rises in an English reading passage. By contrast, in the same context in their L1, they realized 46% of all nuclear tones as rises. Thus, evidence against L1 influence was found.

In one of the very few perception studies on L2 intonation, Mok et al. [19] report results based on an intonation selection

task, which the present study replicates, with 40 speakers of Hong Kong English (HKE). The HKE speakers chose intonation patterns similar to those of the BrE native speakers for some sentence types (i.e. statements, continuation, statement-question/echo, yes-no-questions), but performed worse with tag questions. Mok et al. [19] concluded that HKE speakers have only partial knowledge of the English intonation system due to the absence of equivalent structures in their L1, Cantonese, a tone language. This explanation suggests that learners of English (an intonation language) with an L1 that is also an intonation language will perform better in an intonation pattern selection task than HKE speakers.

In order to test this hypothesis and to learn more about the perception and understanding of English intonation patterns by learners of English in different contexts, we replicated Mok et al.'s [19] study with German-speaking learners of English and conducted an intonation pattern forced-choice selection task. A replication of this study will allow us to compare results across L1 backgrounds and determine how far errors in the choice of appropriate intonation patterns can be explained in terms of L1 transfer or general learning mechanisms. Since Mok et al. [19] found a certain amount of variance in the responses of native BrE speakers, we contrast our results for German-speaking learners with the native speaker data.

3. Method

3.1. Participants

20 learners of English (5 male; age 20-28, mean 24), all L1 speakers of German with no speech or auditory problems, participated in the experiment. Two of the learners indicated that they were brought up bilingually (German/Persian, German/Polish), but were dominant in German. All participants were enrolled in an English literature and linguistics degree at a German university at the time of the study, most of them in their fifth year (five participants in their third semester, one undergoing practical teacher training). The students were quite advanced learners of English, either in their final Bachelor's or Master's semesters. The 13 Master's students had achieved a final grade of 2.1 (2 being an equivalent to 'good' or a 'B'; SD=0.37) in their Bachelor's degree. The participants speak a variety of other foreign languages, mostly French, Spanish, and Italian. 10 of the students assessed their own L2 English proficiency as "nearly native", 9 as "advanced", and one as "intermediate". 6 out of 20 participants had stayed in an English speaking country (3 US, 1 UK, 1 Canada, 1 US, and New Zealand) for a period ranging from one month to one year.

The BrE native speaker control group (previously reported in [19]), consists of 25 participants (3 male). They were all university students aged 19 to 34 (mean 19.3 years; SD: ±1.7). None of them reported speech or hearing problems.

3.2. Data

The participants were presented with a short story (length: 523 words) with different sentence types (e.g. statements, questions, sarcasm, etc.). In total, 28 sentences were selected as test items. As shown in Table 1, five possible English nuclear tones were chosen for this experiment, namely fall, rise, fall-rise, rise-fall, and level. The recordings used in this experiment were read by a female native speaker of BrE (the third author). The same story was then read with the different nuclear tones for all of the 28 target sentences. The nucleus

was on the same syllable in each sample, but was read with different tonal patterns. The expected nuclear tones for each sentence type were determined by the third author's initial reading and standard descriptions of BrE [17, 19].

The present study is based on the results of the intonation selection task described in section 3.3 below. The German participants were additionally recorded reading the story before they completed the selection task, but this study focuses on the perception data only.

Table 1: Number of items of each sentence type.

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 20 participants were first allowed to read the entire story. Subsequently they listened to each of the 28 sentences at a time and were presented (in a PowerPoint slideshow) with five different tonal patterns (labeled A, B, C, D, and E) for each utterance, which appeared automatically and in randomized order, one after the other, once the slideshow was started. In a forced-choice task, the participants were then asked to indicate the tonal pattern that seemed most appropriate to them and note down their selection on an answer sheet. They were allowed to replay each set of tonal patterns as often as they wished (there was no need for participants to store the patterns in short term memory when making their choice). A printed version of the story was presented to participants before the start of the task and was available to them during the entire course of the experiment in order to allow them to retrieve contextual information (which might play a role in the interpretation of sarcasm, in particular).

3.4. Statistics

The data were analyzed with a logistic mixed-effects regression model with RESPONSE (expected/other) as dependent variable. The final model included an interaction between L1 (English/German) and SENTENCE_TYPE/NUCLEAR_TONE as independent variables as well as PARTICIPANT and ITEM as random effects (computed with packages LME4 and LSMEANS in R [20-22]). It should be noted that "accuracy rate" here is used as shorthand for nearness to patterns predicted by the third author's initial reading and standard descriptions of BrE [17, 19].

4. Results

4.1. Overview

Post-hoc Tukey tests reveal that the learners reached a significantly lower accuracy rate (mean 54.1%) than the native

control group (72.6%; $p < 0.0001$, $z = 4.8$). The distribution of the accuracy rates is presented in Table 2.

However, as the interaction between L1 and SENTENCE_TYPE/NUCLEAR_TONE indicates, accuracy rates for the learners were not in all cases significantly worse. In fact, the learners reached accuracy rates similar to those of the control group for 18 out of 28 items, including statement questions, wh-questions, yes/no questions, and echo questions. Moreover, in one sarcasm condition (with a rise-fall tone), the learners significantly outperformed the control group (BrE=8% expected patterns, GE=70% expected patterns; $p < 0.05$, $z = 3.8$). The learners were also more accurate than the native speakers in one of the continuation conditions (BrE=12%, GE=55%). However, this was not the case in the two other continuation conditions, so that overall accuracy for this condition did not differ significantly between the two groups (BrE=25%, GE=30%; n.s., $z = 2.7$).

Table 2: *Distribution of choices for each sentence type.* (rounded percentages; highlighted cells are the expected answers; G=German learners; E=English speakers)

Type	Group	Fall	Rise	Fall-rise	Rise-fall	Level
Statement	G	77	15	5	2	2
	E	96	0	1	1	1
Continuation	G	48	28	18	3	2
	E	57	20	17	0	5
Statement question	G	4	65	28	0	3
	E	6	66	26	0	2
Statement question/echo	G	20	65	10	0	5
	E	24	52	20	4	0
Echo question	G	15	30	45	2	8
	E	9	45	41	1	1
Yes/no-question	G	33	37	22	7	2
	E	36	43	20	1	0
Wh-question	G	78	8	2	3	8
	E	99	0	0	1	0
Closed tag	G	48	25	8	3	15
	E	92	1	1	4	1
Open tag	G	20	55	10	13	3
	E	6	90	0	0	4
Checking tag	G	25	60	0	15	0
	E	4	88	0	4	4
Sarcasm (1)	G	38	10	25	3	25
	E	68	0	2	28	2
Sarcasm (2)	G	15	0	15	70	0
	E	4	0	84	8	0
Checking	G	60	5	25	10	0
	E	56	0	44	0	0

4.2. Similarities between learners and native speakers

Both groups performed most accurately in the case of statement sentences and wh-questions. However, the learners performed marginally worse than the control group, with 77% accuracy for statements (BrE=96%; n.s., $z = 3.0$) and 78% for wh-questions (BrE=99%; n.s., $z = 2.9$). While the native speakers chose almost exclusively falling tones in these conditions, some of the learners chose a rise as a possible contour for statements (GE=15%), and rises and levels for wh-questions (GE=8% for each contour).

The responses of the learners were fairly similar to those of the control group in most conditions, both in overall accuracy rate and in terms of which alternative contours they chose. For continuation sentences, the learners tended to choose mostly falls (48% on average) instead of a rise or level contour, which would have been the expected answers. However, both the average accuracy rate (as mentioned above) and the alternative choices (falls and fall-rise contours instead of rises) are surprisingly similar for both groups.

For the only checking sentence in the experiment, accuracy was insignificantly lower for the learners than for the control group (BrE=44%, GE=25%; $p = 1.0$, $z = 1.3$), but both groups mostly selected a falling tone (BrE=56%, GE=60%).

As far as statement questions are concerned, the accuracy rates for two rise patterns (statement question) and one rise/fall-rise combination (statement question/echo), was again surprisingly similar. As shown in Table 2, the statement questions reached an averaged accuracy of 65% for the 20 non-native speakers and 66% for the native speakers ($p = 1.0$, $z = 0.1$). For the statement question/echo, the learners reached 75% and the native speakers 72% accuracy ($p = 1.0$, $z = 0.2$).

Echo questions exhibit very similar distributions for both groups. The learners reached an accuracy of 75% and the native speakers were 86% accurate ($p = 0.99$, $z = 1.7$). Both groups also accepted simple falls for echo questions.

For yes/no questions, accuracy rates were relatively low for both groups, with 43% accuracy for the BrE speakers and 37% for the German learners ($p = 1.0$, $z = 0.6$). The alternative answers were again quite similar. Instead of a rise contour (expected answer), both groups opted for falls (GE=33%; BrE=36%) and fall-rises (GE=22%; BrE=20%).

4.3. Differences between learners and native speakers

However, there were also several conditions in which the learners deviated from the responses of the control group. These conditions involved tag questions and sentences expressing sarcasm. The greatest differences were observed for closed tag questions, followed by open and checking tag questions. For closed tag questions, the learners chose the expected answers in only 48% of all cases, as opposed to 92% expected answers for the native English group, a highly significant difference ($p < 0.001$, $z = 4.8$). The expected answer here was a fall. Instead, the learners chose a variety of other tones, mostly rise (25%), level (15%), and fall-rise tones (8%).

The learners also performed worse than the native speakers for open tag questions (BrE=90%; GE=55%), but this difference was of only marginal significance ($p = 0.085$; $z = 3.5$). Here, too, the learners opted for a variety of other tones, and opted for falls (20%), rise-falls (13%), and fall-rises (10%), instead of the expected rising contour.

In the case of checking tag questions, the learners were also less accurate than the control group (BrE=80%, GE=60%), but not significantly so (n.s., $z = 2.1$). Instead of choosing a rising contour for this sentence type, some of the learners opted for falls (25%) and rise-falls (15%).

Apart from tag questions, substantial differences occurred in sentences expressing sarcasm. The experiment included two types of conditions involving sarcasm: (1) realized either with a fall or a rise-fall as a nuclear tone, and (2) with a rising-falling nuclear tone. The second type, where the learners outperformed the control group, was discussed above. A contrasting result is revealed by the other sarcasm condition,

where the learners scored significantly below the native group (BrE=96%, GE=41%; $p=0.001$, $z=4.6$). Instead of a rise-fall, most of the learners chose a level or fall-rise (with 25% each).

4.4. Differences between individual speakers

In addition to accuracy rates across the two groups, we also analyzed how uniform the participants were in choosing the expected tones in the experiment. Accuracy rates for most of the British participants ranged from 70% to 80%, and only three participants scored 60% or below. By contrast, only four of the learners reached accuracy rates similar to those of the control group (78.6%, 75.0%, 71.4%, and 67.9%), with the remaining participants achieving between 39% and 57%.

An analysis of the biographies of the four learners that excelled at the perception task reveals that they are likely to have received more (high quality) input than most of the other participants and might also generally be more adept at learning foreign languages than the average participant. In contrast with most of the other learners, they were all enrolled in a Master's program and over 25 of years of age; all speak more than three foreign languages (with one student being a German/Persian bilingual); and three of them had stayed for a longer period (four months to one year) in an English-speaking country. In total, only 6 of the 20 students had spent time abroad and 3 of these 6 students were those that performed best. The other two students with experience abroad did not perform well at this task, but had spent less time abroad (one and six months, respectively). Gender also appears to be a factor correlating with performance in the perception task, with three of the five male participants reaching the lowest accuracy rates overall.

5. Discussion

Based on an intonation pattern selection task, this study investigated whether L2 learners of English possess knowledge of appropriate intonation patterns in different narrative contexts, and, in particular, whether L1 speakers of an intonation language have more accurate knowledge of English intonation patterns than speakers of a tone language [17]. The results reveal that advanced German-speaking learners of English possess appropriate knowledge of some intonation patterns in certain contexts. Both learners and native speakers exhibit similar distributions of accuracy rates and chose similar alternative tones for most sentence types such as continuation sentences, statement questions, statement/echo questions, echo questions, and yes/no-questions revealed almost identical results for both groups. In one particular item involving sarcasm, the learners even outperformed the native speakers. However, a substantial number of the learners appear to lack knowledge of the appropriate intonation patterns in specific contexts. The four sentence types that were the most challenging for the learners involved sarcasm (1) (41%), closed tag (48%), open tag (55%), and checking tag questions (60%).

A comparison of the syntax of English and German reveals that the learners performed well in conditions where the English sentence types have equivalents in the learners' L1. By contrast, tag questions, which have no equivalent in German, were among the most challenging contexts. This suggests that positive L1 influence fosters the L2 acquisition of context-dependent knowledge of intonation (research question 2). However, L1 influence cannot account for the low accuracy for sentences involving sarcasm. An explanation

might be found in the fact that sarcasm is a purely pragmatic phenomenon without syntactic representation.

Given that the learners in this study were L1 speakers of an intonation language, we expected them to perform better at this task than the HKE speakers in [17], whose L1, Cantonese, is a tone language. However, the results for these two groups reveal some similarities. Both encountered problems with tag questions and sarcasm (1). By contrast, the HKE speakers performed worse than the L1 German speakers in the present study in the case of statement questions and wh-questions.

Thus, the hypothesis that speakers of an intonation language would perform better at an English intonation pattern selection task than speakers of a tone language could only be partly confirmed. A more crucial factor appears to be that knowledge of intonation patterns is context-dependent, as both German and Cantonese lack tag questions. Moreover, it might be inconsequential that speakers of German learn English as a foreign language and HKE speakers use English as a second language. In fact, previous research has revealed that the difference between ESL and EFL is one of degree [23-26]. Thus, our results could point towards universal perception errors across L2 groups, but need to be corroborated through the inclusion of further L2 speaker groups.

As for the choice of particular non-target like intonation patterns, our results align with previous research. Instead of falls, a minority of the learners opted for rises (statement sentences and wh-question) and also level tones (for wh-questions). This was also found in the case of tag questions, where many learners chose either a rise or a fall where the opposite would have been the expected choice. This corroborates previous results [16] that state that learners of English replace rises with falls (and vice versa).

Another factor influencing learners' knowledge of intonation appears to be the quality and quantity of input they received. Learners who were successful at the task tended to be more advanced in their studies, and thus are likely to have received a greater quantity of input. They also tended to have spent time in an English-speaking country, and thus have likely received better quality input. While these conclusions are necessarily limited by the size of our sample, they are in keeping with L2 speech learning theories, which stress the role of the quantity and quality of input in advancing the acquisition of L2 phonology [7, 8]. In future research, we will compare our production data with the present results and hope to probe the link between perception and production thought to be so crucial in L2 speech learning theories [7, 8, 27].

Finally, a critical evaluation of our methodology shows that even the native speakers did not perform at ceiling level in all conditions. This might either suggest that, even for native speakers, this is a challenging task, or that our notion of what is the expected choice of an intonation pattern in these contexts needs to be re-evaluated. In any case, our approach, relying on an explicit comparison of a group of learners with a native control group, ensured that task-specific effects can be distinguished from effects reflecting the learners' (lack of) knowledge of the English intonation system.

6. Acknowledgements

The authors would like to acknowledge the assistance of Dr Noor Mat Nayan in the collection of the UK data.

7. References

- [1] 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.
- [2] 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.
- [3] 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.
- [4] T. Scovel, "A critical review of the critical period research," *Annual Review of Applied Linguistics* 20, pp. 213-223, 2000.
- [5] E. Grabe, B.S. Rosner, J.E. Garcia-Albea, and X. Zhou, "Perception of English Intonation by English, Spanish and Chinese listeners," *Language and Speech*, 46, pp. 375-401, 2003.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] S.M. Gass and L. Selinker, *Second Language Acquisition: An Introductory Course*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates, 2001.
- [10] 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.
- [11] P. Roach, *English Phonetics and Phonology: A practical course*. Cambridge: Cambridge University Press, 2009.
- [12] M. Wrembel, "Metacompetence-based approach to the teaching of L2 prosody: practical implications," *Non-Native Prosody – phonetic description and teaching practice*, eds. J. Trouvain & U. Gut. Berlin: Mouton de Gruyter, pp. 189-209, 2007.
- [13] J. Esser, "Contrastive intonation of German and English," *Phonetica* 35, 41-55, 1978.
- [14] A. Fox, *German Intonation: An Outline*. Oxford. Clarendon Press, 1984.
- [15] 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.
- [16] U. Gut, *Non-native speech: a corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt: Lang, 2009.
- [17] A. Cruttenden, *Gimson's Pronunciation of English (8th edition)*. London: Routledge, 2014.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] D. Bates, M. Maechler, B. Bolker and S. Walker. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, <URL: <http://CRAN.R-project.org/package=lme4>>, 2014.
- [22] R. V. Lenth and M. Hervé. lsmeans: Least-Squares Means. R package version 2.17. <http://CRAN.R-project.org/package=lsmeans>, 2015.
- [23] A. Edwards. "The EFL–ESL Continuum and the Case of the Netherlands: A Comparative Analysis of the Progressive Aspect." *World Englishes* 33: 173–94, 2014.
- [24] R. Fuchs and E.-M. Wunder. "A sonority-based account of speech rhythm in Chinese learners of English," in *Universal or diverse paths to English phonology? Bridging the gap between research on phonological acquisition of English as a second, third or foreign language*, U. Gut, R. Fuchs & E.-M. Wunder, pp. 165–184. Berlin: de Gruyter., 2015.
- [25] U. Gut, R. Fuchs and E.M. Wunder, eds., "Universal or diverse paths to English phonology." Berlin: de Gruyter, 2015.
- [26] G. Gilquin, "At the interface of contact linguistics and second language acquisition research: New Englishes and Learner Englishes compared." *English World-Wide* 36(1), 91-124, 2015.
- [27] 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.