Prosody in Turkish learners of German as a Foreign Language

Sabine Zerbian [1], *Jane Kühn* [2], *Christoph Schroeder* [2], *Svenja Schuermann* [2]

[1] Institute of Linguistics: English, University of Stuttgart, Germany
[2] SFB 632 "Information Structure", University of Potsdam, Germany
sabine.zerbian@ifla.uni-stuttgart.de, {jkuehn,schroedc}@uni-potsdam.de

10.21437/SpeechProsody.2014-41

Abstract

Results of a pilot study are presented which investigates the prosodic realization of information structure by six learners of German as a Foreign Language (GFL) with Turkish as first language. Question-answer pairs were read out loud, which systematically varied the position of narrow focus in the response by means of a preceding *wh*-question.
A qualitative analysis of the results shows deaccentuation of postfocal constituents in the case of subject focus for 4/6 GFL-speakers but no consistent pitch increase on focused constituents. Two speakers did not change prosody due to information structure.
The results are discussed in connection with the acquisition of prosody as a marker of information structure. Deaccentuation has been reported to cause problems in L2 prosody. In Turkish, deaccentuation occurs postfocally. The claim will be motivated that the occurrence of deaccentuation in the L1 is a necessary but not sufficient condition for early acquisition of deaccentuation in a foreign language.

Index Terms: L2 prosody, German, Turkish, focus, deaccentuation, production

1. Introduction

The acquisition of L2 prosody, i.e. linguistically relevant changes of fundamental frequency (F0), intensity and/or duration over the course of an utterance, has received comparatively limited attention in the literature on Second or Foreign Language Acquisition. According to [1:57] and referring back to the intonational typology in [2], phonological influences in L2 prosody must be differentiated from phonetic influences. Phonological influences on L2 prosody result from differences in the inventory of phonological tunes, their form, and in the meanings assigned to the tunes. Phonetic influences result from a difference in the phonetic realization of an identical phonological tune.
Previous studies on L2 of non-related languages have shown that the target-like placement of sentence accent (phonological level) as well as the target-like realization of pitch accents (phonetic level) cause difficulties in foreign language acquisition. One of the particular difficult features reported in studies like e.g. [3-7] lies in the function of prosody in languages like English and German, where sentence-level prosody indicates information structural notions such as topic, focus and givenness. For L2 prosody, a non-target-like reaccentuation of constituents that are given through the preceding discourse has repeatedly been reported in the studies cited above.
The article reports the results of a pilot study into the prosody of Turkish learners of German as a foreign language. The functional use of prosody as marking focused and given constituents will be central. The German-Turkish language pairing is interesting, because German, just like English, shows both prosodic focus marking through an increase in fundamental frequency (F0) and duration, as well as deaccentuation of given constituents. For English, prosodic focus marking and deaccentuation have been considered "opposite sides of the same coin" [8:67]. Turkish is not related to German. However, it can be considered to have a similar prosodic representation, being a stress language which also uses prosody at the sentence level. However, it shows less systematic prosodic focus marking (see discussion in section 2.2) but has postfocal deaccentuation[1].
The article is structured as follows: Section 2 gives the relevant background on the prosody of the two languages involved. Section 3 presents the experiment with a summary of the results in section 3.5. Section 4 provides a discussion.

2. Focus prosody in German and Turkish

2.1. German

German has been classified as an intonation-only language (cf. [9]) which has lexical stress. Postlexically, it uses different types of pitch accents together with boundary tones in order to express pragmatic contrasts (see [10]). The interaction of information structure and intonation is uncontroversial for German.
Default sentence accent (in all-new sentences) is assigned on the basis of syntactic structure: every argument of the verb is accented, and the verb might also be accented if there is phrasal integration of the verb and its immediately preceding argument. Perceptually, the last accent in a sentence is perceived to be the most prominent. In terms of phonetic implementation of the accents, downstepping of high-toned accents has been reported to occur, i.e. realization of a high tone at a lower phonetic level relative to a preceding high tone [11].
Under narrow focus, the most prominent accent occurs on the focused constituent. As experimental work by [11] for German confirms, pitch is raised under focus when compared to a baseline of all-new sentences. Givenness leads to a lowering of pitch, whereby the position of the given constituent with respect to the focus is relevant. If the given constituent occurs preceding the focus of the sentence (prenuclear position), the constituent will have a pitch accent with a comparatively low pitch. If the given constituent occurs following the focus of the sentence (postfocal position), it has been said to be deaccented. This overall pattern was confirmed by German native speaker controls, participating in the same experimental task as reported in the current article.

[1] The difference between deaccentuation and postfocal compression is not always clear-cut in actual data. The former refers to the deletion of a pitch accent whereas the latter refers to the compression of the pitch range of a pitch accent. Here, the term deaccentuation is used, bearing in mind that this issue requires further investigation.

Next to F0, increased duration on the focused constituent is a further cue to focus in German with given constituents only being shortened in prefocal position [12].

2.2. Turkish

Turkish is an SOV-language. Though differing in details of analysis, scholars agree that the default sentence accent occurs on the last argument before the verb, which is normally the object [13, 14]. In addition, an accent can occur on every argument.

As for the expression of focus, [15] differentiates between a syntactic strategy in which the preverbal position is used for information focus (e.g. question-answer pairs) where the answers is not accessible from context. In cases of e.g. correction focus where answer can be chosen from an explicit set of alternatives, a constituent is marked as focused in situ by means of a pitch accent. Along the same lines, [16] describes that for both contrastive and non-contrastive emphasis prosodic means such as strong stress and high pitch are used. However, in addition to these prosodic means, non-contrastive emphasis is further marked by placing the focused constituent in the immediately preverbal position. [17], on the other hand, do not assign the immediately preverbal position a primary role in focus assignment. In their examples, focused constituents always bear stress and can occur in any preverbal position.

To our knowledge, only one acoustic study on focus in Turkish exists. [18], in using a comparable methodology as outlined in the present paper, finds that there is no increase in F0 on the focused constituent in the Turkish data she collected. She found rather unsystematic significant differences for the other acoustic correlates: A sentence-initial focused subject differs acoustically from its neutral counterpart in terms of longer duration, a sentence-final verb differs in being higher in intensity. An object did not show differences between the focused condition and the neutral counterpart in any of the three acoustic measures F0, intensity and duration. Unexpectedly, an F0 increase on the preceding object is found when the final verb is in focus. For givenness marking, [18] found that in cases of subject focus there is a significant F0 drop in the post-focus domain. The same pattern does not hold for object focus though. As tested by means of a perception experiment, subject focus has the highest recognition rate, probably due to the salient F0 drop.

It is uncontroversial that in Turkish sentence accent never occurs in the postverbal domain. The prosody in the postverbal domain has been described as being without "any indicator of prosodic structure" [19:144]. [19: fig. 3] provide a pitch track showing that we can expect to find deaccentuation postfocally, not only postverbally. In their example, the pitch contour remains entirely flat following the constituent in focus, which in this specific case is a sentence-initial subject. Their observation is thus in line with the results on the F0 drop following a focused subject in [18].

2.3. Hypotheses

The research question is whether Turkish learners of German change prosody depending on information structure and if so, by means of which acoustic cues. More precisely:

- Do Turkish learners of German use increased F0 on the focused constituent? Both languages seem to differ on the acoustic cues used for prosodic focus marking: Whereas German relies on F0 as well as on duration, Turkish does not seem to rely on F0 as an acoustic cue to prosodic focus.
- Do Turkish learners of German use deaccentuation to signal givenness? Although both languages use deaccentuation for postfocal constituents, research on the acquisition of L2 prosody has reported that learners have difficulties with this feature ([3-7]).

3. Experiment

3.1. Task

An elicited-production study was conducted using the same methodology as in [20] and [18]. German target sentences were elicited that differed in the focused constituent only. The target sentences consisted of simple SVO-structures, and the focused constituent was either the subject, the verb or the object. As a baseline condition, a broad focus rendering was elicited first.

The target sentences were presented in writing, together with a preceding question unambiguously eliciting the desired focus structure, and accompanied by a picture illustrating the action. In addition, the focused target words were underlined in order to reduce errors (cf. [20]). This experimental setup has proven in various studies to be successful at eliciting the predicted focus structures (cf. [20] and subsequent studies).

3.2. Target sentences

Target sentences were constructed in such a way as to systematically control a number of factors and to be comparable across the five different target sentences. To this end, all sentences displayed the same number of words. To minimize segmental variation, the target words showed the same stress pattern (initial stress), the same phonological length in the stressed vowel, and the same segmental make-up (CV with sonorant consonants wherever possible). The object nouns were disyllabic to allow for pitch accents and boundary tones to be realized without tonal crowding.

In constructing five different target sentences as shown in (1), we opted for type repetitions instead of token repetitions in order to increase the diversity of the targets.

(1) Target sentences
   a. Lena malt ein Lama.
      PROP.NAME paint ART llama
      'Lena is painting a llama.'
   b. Nele wohnt in Meißen
      PROP.NAME live PREP PROP.NAME
      'Nele lives in Meißen.'
   c. Nina webt das Leinen.
      PROP.NAME weave ART linen
      'Nina is weaving the linen.'
   d. Heiner baut die Mühle.
      PROP.NAME build ART mill
      'Heiner builds the mill.'
   e. Maya liest die Bücher.
      PROP.NAME read ART books
      'Maya is reading the books.'

3.3. Participants

Nine Turkish learners of German participated in the study. Participants were students of Translation Studies in German at Ege University in Izmir, where the recording took place in

September 2012. The data were collected by the second author, communicating with the participants in German. The data of one speaker (speaker 1) was not recorded properly due to technical problems, the data of two further speakers (sapeakers 2 and 7) had to be excluded because only two target sentences were recorded.
Of the remaining 6 speakers, 5 were female and one was male. The participants were between 20 and 25 years old. They all spoke Turkish as an L1, had good to very good knowledge of English (self-reported) and had studied German as a Foreign Language (GFL) for 3 to 6 years, including a preparatory language course at the University which has been designed to lead to level B1/B2 (Common European Framework of Reference for Language Teaching (CEFR)).

3.4. Analysis

The analysis concentrates on fundamental frequency (F0) as the main correlate of sentence prosody, and more specifically prosodic focus marking and deaccentuation in the target language German. All target sentences were segmented into syllables and 10 measurements of F0 were taken for each syllable, using ProsodyPro [21]. Data were checked manually for spurious pitch values. The datapoints were averaged across the five utterances of the same focus condition for each speaker, and are presented in figure 1 (see next page) for each speaker individually (in the same pitch range of 50-250 Hz, maintaining individual differences in F0 expansion).
In addition, a ToBI annotation of the averaged pitch contours is provided in table 1. Tone labels were assigned following these conventions: H* for each discernible pitch peak, H*+L for a pitch peak with a clear fall on the same constituent, !H* if the pitch peak is lower than the preceding pitch peak, ↑H* for a pitch peak that is higher than the preceding pitch peak. In some cases (e.g. speaker 9), labels had to be assigned based on deviations from expected pitch transitions if there had been no tone target. No evidence was found for boundary tones at intermediate phrases. However, the ToBI annotation should be seen as a phonetic rather than phonemic annotation of the produced intonation, given that further research needs to be done for the phonemic status of the pitch accents transcribed.

Table 1: ToBI annotation based on the averaged pitch tracks

speaker	focus	S	V	O
3	broad	H*+L	!H*+L	↑H*+L L%
	subject	H*+L		L%
	verb		H*	↑H*+L L%
	object	H*+L	!H*	↑H*+L L%
4	broad	H*+L	↑H*	L%
	subject	H*+L		L%
	verb	H*+L	↑H*	L%
	object	H*+L	↑H*	L%
5	broad		H*+L	!H* L%
	subject	H*+L		L%
	verb		H*+L	L%
	object		H*+L	!H*+L L%
6	broad		H*+L	!H* L%
	subject		H*+L	!H* L%
	verb		H*+L	H* L%
	object		H*	H* L%
8	broad		L+H*	!H* L%
	subject		L+H*	!H* L%
	verb		L+H*	!H* L%
	object		L+H*	!H* L%
9	broad	L*H	!H*	!H*+L H%
	subject	L*H		!H*+L H%
	verb	L*H	L+H*	LH%
	object	L*H	H*	!H*+L H%

3.5. Results

The observations are given for each speaker individually first, before they are summarized (number refers to speaker):

- 3: on-focus increase in F0, deaccentuation only in case of subject focus
- 4: no on-focus increase in F0, deaccentuation only in subject focus
- 5: on-focus increase in F0, deaccentuation in subject and verb focus
- 6: no on-focus increase in F0, despite a lower F0 in subject focus, the F0 patterns are qualitatively the same across all focus conditions
- 8: identical F0 contours across all focus conditions
- 9: on-focus increase in F0 in verb focus only, deaccentuation only in subject focus

All speakers realize intonation in broad focus context similar to object focus. Beyond that, as can be expected in a group of learners, we find considerable heterogeneity in the data. The speakers can be described to fall into roughly two groups: speakers 6 and possibly 8 do not show any change in the F0 pattern dependent on the focus structures of the sentences uttered. The other speakers do show a change in prosody related to information structure, as detailed above.
To sum up, of the speakers investigated, most (though not all) use some kind of prosody in connection with information structure (cf. discussion section). Deaccentuation is the most reliable cue used by four out of six speakers, but only in the case of subject focus. Half the speakers also use an increase in F0 on focused constituents. It needs to be noted though, that only speaker 3 links the overall F0 pattern to information structure in a way comparable to the target language.
Given that [18] reported duration and intensity as cues to focus in Turkish rather than F0, these parameters were checked quantitatively for those two speakers who did not show any F0 changes due to information structure (speakers 6 and 8). Only for speaker 6 we find higher duration and intensity on a focused constituent, when compared to the broad focus condition, although the differences are not significant.

4. Discussion

The production data show that some Turkish learners of German do not change prosody according to information structure, neither using phonetic cues of the L1 like intensity and duration [18], nor phonetic cues of the L2 such as F0 increase on focused constituents nor phonological features shared by L1 and L2 such as deaccentuation on given constituents.
Others learners, however, do show a change of prosody due to information structure. Interestingly, these speakers have in common that they show deaccentuation in the case of subject focus. Postfocal deaccentuation in the case of subject focus is a prosodic feature shared by the L1 Turkish and the L2 German and it is a very prominent cue to a shift of focus to the subject ([18] for perception results, auditory impression suggests the same for our data). Some of those speakers also show an increase of F0 on the focused constituent when compared to the broad focus condition and/or deaccentuation on given constituents other than the subject.
Two results thus need further discussion: First, the observation that some learners produce postfocal deaccentuation in cases

of subject focus, although deaccentuation has been reported as a difficult feature to be acquired in foreign language acquisition. Second, the observation that some learners do not change prosody at all under the influence of information structure.

The studies on L2 prosody which constitute the claim that deaccentuation is a difficult feature, investigate language pairings with English as L2 and L1s which either show equivalents to deaccentuation in English but whose prosodic systems are typologically very different (e.g. English, Korean and Japanese in [4] or English, Mandarin and Korean in [7]) or which do not have deaccentuation themselves [5, 22, 23]. [24] have argued that phonological representation is crucial in foreign language acquisition in that representational differences between two languages cause difficulties. If transferred to the suprasegmental domain, Korean and Japanese learners of English are predicted to have difficulties with English prosody due to the representational differences between the two prosodic systems involved, independent of any specific function of the target prosody.

[5, 22, 23] investigate the use of prosody by Spanish learners of English. Spanish and English can both be classified as stress and intonation-only languages and thus share a similar prosodic representation. Postfocal deaccentuation was still found to be difficult for these learners. Here we need to note that L1 Spanish is said to not use deaccentuation (e.g. [25]).

Prosody used for information structure is a typologically marked feature [26]. Following the Markedness Differential Hypothesis [27], this makes it a difficult feature to be acquired, which has been confirmed by the studies on Spanish learners of English. However, in the case of Turkish learners of German, the L1 shares the marked feature as discussed in section 2.2. This might be the reason why even learners at the level of B1/B2 show the use of this feature. The lack of prosodic marking of focused constituent by a gradient increase of F0 on these constituents can be interpreted along the same lines. Prosody used for information structure is a marked feature whose acquisition is difficult, especially if the feature is not shared with the L1. As Turkish does not use F0 increase on focused constituents, this new feature has been acquired by very few learners only.

It can thus be argued that the presence of a marked feature in the L1, such as postfocal deaccentuation, is a necessary condition for the acquisition of that feature in the L2. The results of our production study equally show that it is by no means a sufficient condition. Two speakers do not show any changes in prosody due to information structure, not even in those acoustic correlates which are allegedly used in Turkish (intensity and duration according to [18]). How can the lack of any prosody in speech production be accounted for? We follow [28] and [7], who observe that under certain circumstances learners seem to fall back on some default intonation patterns. [28] investigated the prosody of advanced German learners of English and speculated that observable errors in the placement of pitch accents give "room for the assumption that non-native speakers develop a consistently simplified assignment of tonal categories, a kind of default accent" [28: 87]. This would be in line with [29]'s "basic variety" and a corresponding reduced prosodic realization of the target language. In some of her experimental tasks, [7] observes that Mandarin and Korean learners of English rely on common prosodic patterns used for broad focus.

We thus interpret the results of our pilot study as evidence for the following developmental path in the acquisition of IS-related intonation by Turkish learners of German:

1. Default intonation: One all-purpose default pattern, not differentiating e.g. different kinds of narrow focus.
2. Transfer of L1 feature: Transfer leads to postfocal deaccentuation with subject focus. This is possible because L1 Turkish shares this marked feature with German.
3. Acquisition of a new phonetic feature, namely the gradient increase of F0 on the focused constituent, shown by only very few learners of our group.

Further research is clearly needed though in order to confirm the results.

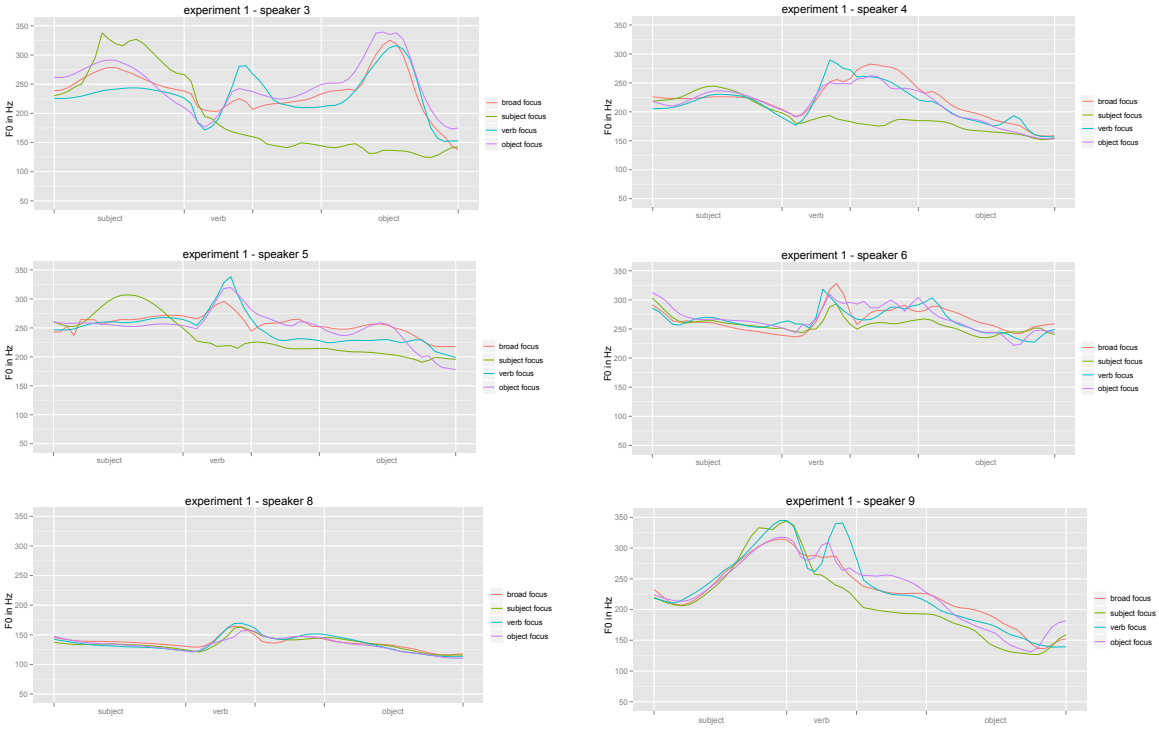

Fig. 1: averaged F0 contours for the four focus conditions for each speaker individually (pitch range of 50-250 Hz)

5. References

[1] Mennen, I., “Phonological and phonetic influences in non-native intonation”, in J. Trouvain & U. Gut [Eds], Non-native Prosody: Phonetic Descriptions and Teaching Practice, 53-76, Mouton de Gruyter, 2007.

[2] Ladd, D.R., “Intonational Phonology”, Cambridge University Press, 1996.

[3] McGory, J., “The Acquisition of Intonation Patterns in English by Native Speakers of Korean and Mandarin”. PhD dissertation, Ohio State University, Columbus, 1997.

[4] Ueyama, M., Jun, S.-A., “Focus Realization in Japanese English and Korean English Intonation”, Japanese and Korean Linguistics 7:629–645, 1998.

[5] Verdugo, D.R., “Prosodic realization of focus in the discourse of Spanish learners and English native speakers”, World Englishes 14, 9–32, 2006.

[6] Gut, U., “Non-native Speech. A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German”, Peter Lang, 2009.

[7] Baker, R.E., “The Acquisition of English Focus Marking by Non-Native Speakers”, PhD dissertation, Northwestern University, Evanston, Illinois, 2010.

[8] Ladd, D.R., “The Structure of Intonational Meaning - Evidence from English”, Indiana University Press, 1980.

[9] Gussenhoven, C., “The Phonology of Tone and Intonation”, Cambridge University Press, 2004.

[10] Féry, C., “German Intonational Patterns“, Max Niemeyer, 1993.

[11] Fery, C., Kügler, F., “Pitch accent scaling on given, new and focused constituents in German”, Journal of Phonetics 36:680–703, 2008.

[12] Kügler, F., “The role of duration as a phonetic correlate of focus”, in P. Barbosa, S. Madureira & C. Reis [Eds], Proceedings of Speech Prosody. Editora RG/CNPq, 591–594, 2008.

[13] Kan, S., “Prosodic Domains and the Syntax-Prosody Mapping in Turkish”, Master's Thesis, Bogazici University, Bogazici, 2009.

[14] Kamali, B., “Topics at the PF Interface of Turkish”, PhD dissertation, Harvard University, Cambridge, Massachusetts, 2011.

[15] Issever, S., “Information structure in Turkish - the word order-prosody interface”, Lingua 113:1025–1053, 2003.

[16] Kornfilt, J., “Turkish Grammar”, Routledge, 1997.

[17] Göksel, A., Özsoy, A.S., “Is there a focus position in Turkish?”, in A. Göksel & C. Kerslake [Eds], Studies on Turkish and Turkic Languages. Proceedings of the Ninth International Conference on Turkish Linguistics, Harrassowitz, 219–228, 2000.

[18] Ipek, C., “Phonetic realization of focus with no on-focus pitch range expansion in Turkish”, in Proceedings of the 17th International Phonetic Congress (ICPhS), 140–143, 2011.

[19] Özge, U, Bozsahin, C., “Intonation in the grammar of Turkish”, Lingua 120:132–175, 2010.

[20] Xu, Y., “Effects of tone and focus on the formation and alignment of f0 contours”, Journal of Phonetics 27:55–105, 1999.

[21] Xu, Y., “ProsodyPro - A Tool for Large-scale Systematic Prosody Analysis”, in Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013). Aix-en-Provence, 7–10, 2013.

[22] Nava, E., “Prosody in L2 acquisition”, in 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007). Cascadilla Proceedings Project, 2007.

[23] Nava, E.; Zubizaretta, M.L, “Deconstructing the Nuclear Stress Algorithm: Evidence from second language speech”, in N. Erteschik-Shir & L. Rochman [Eds], The Sound Patterns of Syntax. Oxford University Press, 2010.

[24] Eckman, F.R., Elreyes, A., Iverson, G.K. “Some principles of second language phonology”, Second Language Research 19 (3): 169–208, 2003.

[25] Cruttenden, A. “The de-accenting of given information. A cognitive universal?”, in G. Bernini & M.L. Schwartz [Eds], Pragmatic Organization of Discourse in the Languages of Europe. Mouton de Gruyter, 311–355, 2006.

[26] Rasier, L., Hiligsmann, P. “Prosodic transfer from L1 to L2. Theoretical and methodological issues”, Nouveaux cahiers de linguistique francaise 28: 41–66, 2007.

[27] Eckman, F. “Markedness and the Contrastive Analysis Hypothesis”, Language Learning 27 (2): 315–330, 1977.

[28] Jilka, M., “The contribution of intonation to the perception of foreign accent. Identifying intonational deviations by means of F0 generation and resynthesis”, PhD thesis, University of Stuttgart, 2000.

[29] Klein, W., Perdue, C. “The Basic Variety (or: Couldn't natural languages be much simpler?)”, Second Language Research 13 (4): 301–347, 1997.