TESTING THE CONTRIBUTION OF PROSODY TO THE PERCEPTION OF FOREIGN ACCENT

Matthias Jilka
University of Stuttgart

Abstract:

As of yet the contribution of prosodic factors, particularly intonation, to the perception of foreign accent has seldom been the subject of studies in second language acquisition. In perception tests using low-pass filtered stimuli this study shows that a significant number of listeners is able to discriminate American English and German in stretches of speech 30 seconds or longer. The listeners are even more successful in direct pair comparisons, showing that the prosodic characteristics of American English and German are distinct enough to contribute to an obvious language-specific difference. The same tests are repeated using low-pass filtered stimuli with monotonous intonation. In this case all results are significantly worse. It can thus be claimed that intonation is clearly the most important prosodic factor in the perception of foreign accent. The relevance of intonation is also validated in a foreign accent rating task and a direct pair comparison using original foreign-accented American English utterances by native speakers of German and F_0-generated and resynthesized versions in which the intonational errors are corrected. Finally, the study also attempts to give an impression of the respective shares of segmental and intonational aspects of foreign accent by comparing the ratings of fully synthesized versions of the speakers' original utterances (no segmental foreign accent, but intonational foreign accent) and of F_0-generated, "improved" versions (no intonational, but segmental foreign accent). It is shown that segmental aspects have a bigger influence on the perception of foreign accent.

1. Introduction

'Testing the contribution of prosody to the perception of foreign accent' seems to be a very general and broad subject for a study in the field of second language acquisition (SLA), but as a matter of fact there has not yet been a lot of research in this particular area. Studies of either foreign accent, language acquisition or with a speech technological background (automatic speech recognition or speech synthesis) at best touch on general prosodic phenomena or concentrate on closely defined specific aspects.

Contrastive studies do also deal with prosodic differences between languages, but do not necessarily say anything about the production and perception of foreign accent, see e.g., Grabe's description of compression and truncation of phrase-final falling pitch accents by native speakers of British English and German (Grabe 1997, 1998).

On the whole it can be stated that so far there have not been any studies that attempt to give a systematic and coherent account of prosodic aspects of foreign accent with respect to the prevailing models of SLA such as Flege's Speech Learning Model (Flege 1995) or Best's Perceptual Assimilation Model (Best 1995).

Described in more detail the overall aim of this study is to offer a perceptual analysis of the contribution of prosody in general and intonation in particular to the perception of foreign accent, both in absolute terms and in relation to segmental aspects. The latter may be expected to be more prominent, at least in the examination of foreign accent in the American English productions of native speakers of German that is given in this study. As far as intonation is concerned the basis for the perception experiments is provided by a thorough analysis and identification of the intonational deviations detected in the productions of the German speakers. Detailed accounts of this analysis can be found in Jilka (2000). A short summary follows in section 2.

2. Analysis of intonational deviations

2.1. Speech data

Data of foreign-accented speech was collected both for native speakers of German speaking American English and native speakers of American English speaking German. Ten native speakers of the two respective languages, none of whom were paid for their contribution, participated. The American subjects were all aged between 22 and 35, had an accademic background in German literature and language, and had been staying in Germany for at least 6 months at the time of the recordings. Their German speech did not exhibit a conspicuous segmental foreign accent or grammatical mistakes. The German subjects were in the same age group, had a similar academic background in English and experience from longer stays in the United States. Recordings were made in the anechoic chamber of the Institute of Natural Language Processing on digital audio tape (DAT) with a sampling frequency of 44.1 KHz and 16 bit accuracy and took about 30 minutes per session and subject.

The subjects were asked to produce three types of speech: read, prompted (repeated) and unprompted (quasi-spontaneous).

To provide for recordings of read speech the American speakers read the two short stories "Das Dicke Kind" and "Die Buttergeschichte", while their German counterparts read longer news reports from the Boston Radio News Corpus,

in short BRNC (Ostendorf et al. 1995). The subjects were also required to repeat short to middle-length utterances that were played to them. They were specifically not asked to “imitate” what they heard, so as to make the task seem relatively informal and minimize effects of immediate imitation that would prevent conclusions about the speaker’s linguistic performance (see e.g., Markham 1997, pp. 39 - 46). The utterances to be repeated were again taken from the BRNC as well as the ToBI training corpus (Beckman and Ayers 1994) for the German speakers, and from Verbmobil (Wahlster 1997) , the Kiel Corpus (Kohler 1994) and the IMS German Radio News Corpus (Rapp 1998) for the Americans. During the recording sessions the subjects also produced a lot of speech in conversation with the instructor. The conversations took place in an informal atmosphere and concerned a variety of non-academic subjects. Despite the unusual surroundings their productions can be considered spontaneous or at least unprompted.

2.2. Categorical representation

It is important to be aware of the fact that the results of a study of intonational foreign accent, i.e., the identified individual instances of intonational foreign accent, are strongly influenced, even predetermined by the system of transcription that is used to interpret the continuous intonation contour. In other words the theory of intonation representation within which we examine the intonation contour shapes our perception of its relevant units. The chosen framework of intonation description must thus fulfill the specific requirements of the investigation which demand both a category-oriented view of intonational stucture, so that the description is compatible with the leading models of second language acquisition and the possibility of F_0-generation on the basis of the accompanying system of intonation description.

A tone sequence-style approach with ToBI (short for “Tone and Break Indices”) as the corresponding transcription system (see Silverman et al., 1992) was deemed to deal with these criteria best.

The Tone Sequence Model (Pierrehumbert 1980) analyzes the F_0 contour as a series of target values connected by transitional functions. It is primarily meant to be an approach to intonation generation, but also provides the corresponding basic principles of an intonational description.The targets are defined as either high (H) or low (L) on an abstract top- and baseline that encompass a speaker’s local pitch range. The excursion of a specific F_0 pattern can thus be controlled depending on the respective concrete values for top- and baseline, supporting the view that the same pitch patterns, i.e. the same tonal event, may take on different phonetic realizations in different pitch ranges or otherwise variable contexts (e.g. depending on segmental context or speaking rate). The association of one or several targets with stressed syllables or the phrase margins marks them as particular tonal events (pitch accents and boundary tones) that make up the intonation contour. The ToBI system of intonation transcription interprets these tonal events as phonological intonation categories, assigning specific tone labels to them. F_0 generation as it is applied in this study (see Jilka et al. (1999) for a detailed description), is based on ToBI labels which stand for the tonal categories that, like segmental phonological categories, may have different phonetic realizations, the phrasal, tonal and segmental environment determining the precise placement of the target points. The targets are defined in the dimensions of pitch range (top- and baseline) and time (position in the labelled syllable) and connected by means of linear interpolation. The generation process is summarized in Figure 1.

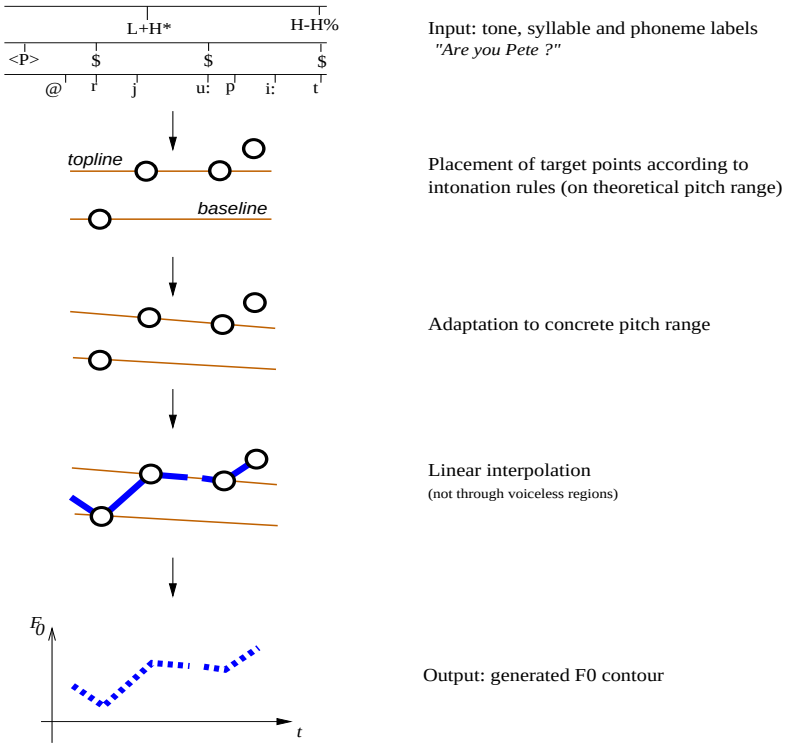

Figure 1. Summary of the F_0 generation process

The process of identifying instances of intonational foreign accent is carried out by means of an analysis-by-synthesis approach: intonational deviations from the postulated norm (ToBI tone label inventories and intonation rules from the generation mechanism) or comparable equivalent utterances by native speakers are examined as to the relevance of their contribution to the impression of foreign accent. The confirmation of their relevance is achieved via the generation of an alternative version of the corresponding original utterance that is completely identical with the original except for the region that (according to the preceding analysis) is assumed to be responsible for the foreign accent. If in a direct comparison the generated/resynthesized version clearly sounds less foreign-accented or even does not show any foreign accent at all, this is taken as confirmation that the observed intonational pattern was indeed the cause for the impression of foreign accent.

The approach is demonstrated by an example in Figure 2, which depicts the utterance “I’m 31 years old” by a native speaker of German while introducing herself at the beginning of the recording session. As can be seen in the top contour the speaker uses an unusual final tune in her production by stressing the word “years” with a rising pitch accent (L+H*) and maintaining that high level (transcribed by a ‘plateau’ H-L% boundary tone) until the end of the phrase. The resulting contour is difficult to interpret because the rise on “years” and the high continuation on “old” seem to put a focus on these two elements that suggests a set of alternative such as “months young.” This, however, makes no sense with respect to the circumstances under which the utterance was made. For the generation of an “improved” version the context allows two possibilities: either a declarative statement or an utterance ending in a continuation rise. The declarative contour with the characteristic L-L% boundary tone (middle contour) would give the utterance a sense of conclusion, while a continuation rise with the corresponding L-H% boundary configuration (bottom contour) would indicate that the speaker will go on to say more about herself (“I’m 31 years old. I work at the University of Stuttgart ...”) as was the case in the actual situation. Both resynthesized versions decidedly improve the speaker’s utterance, showing that the final tune in her original version did indeed contribute to the impression of foreignness.

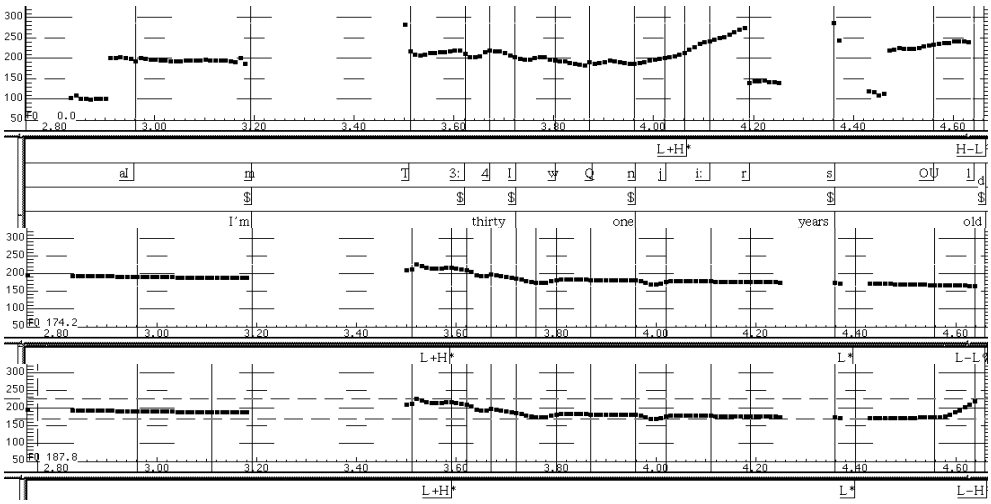

Figure 2. F_0 contour of the phrase “I’m 31 years old” uttered by a native speaker of German. Top: original contour; Middle: generated declarative contour; Bottom: generated continuation rise.

2.3. Assumptions about intonational deviations

In the investigation of the intonational aspects of foreign accent, one has to be aware of the potential problems in identifying instances of intonational foreign accent. In order to facilitate a meaningful ana-lysis a number of assumptions has to be made about intonation in general and its consequences for the study of intonational foreign accent. The aim is to define which intonational phenomenon constitutes a case of foreign accent and which does not.

The definition of what constitutes a case of intonational foreign accent seems fairly straightforward: the intonation in the speech of a non-native speaker must deviate to an extent that is clearly inappropriate for what is considered native. The decision of what intonation is inappropriate or even impossible strongly depends on the surrounding context, much more so than it is the case for deviations in segmental articulation. For this reason the following important assumptions about the significance of variation in intonation seem to be necessary to facilitate a reliable identification of intonational foreign accent.

- intonational deviations can only be associated with foreign accent, if the context allows no other interpretation that would make the complete utterance still acceptable in the L2
- context-dependent variation in intonational realizations is much more frequent than in segments
- foreign accent may be present not in a particular type of intonational variation, but only in the degree and/or frequency of the variation (as native speakers also deviate from the prototypical realization of a tonal category)

The latter assumption implies that slight intonational deviations, that alone or in small numbers (in short utterances) would not be perceived at all, can accumulate in longer stretches of speech and trigger an effect of foreign accent.

2.4. Types of intonational foreign accent

Four major causes of intonational foreign accent were identified. They involve a general incorrect choice and/or placement of tonal categories, the transfer of tonal categories from the speaker's L1 in corresponding discourse situations, transfer in the phonetic realization of tonal categories, and overall characteristics such as relatively more tonal movement in the productions of the American speakers. Most of these manifestations of intonational foreign accent are present in an American speaker's version of the German utterance "... denn man hatte dort am Abend vorher auf einem Schild schon lesen können, daß frische Butter eingetroffen sei" (... because one had been able to read on a sign there already the evening before that fresh butter had arrived), which is depicted in the top contour of Figure 3. The middle contour shows the corresponding version of a native speaker of German, while in the bottom contour the rule-generated, improved, version of the American speaker's original can be seen. The three versions of the example utterance (as well as of Figure 2) are available for auditory evaluation by the reader at http://www.ims.uni-stuttgart.de/phonetik/matthias/sound4diss.htm

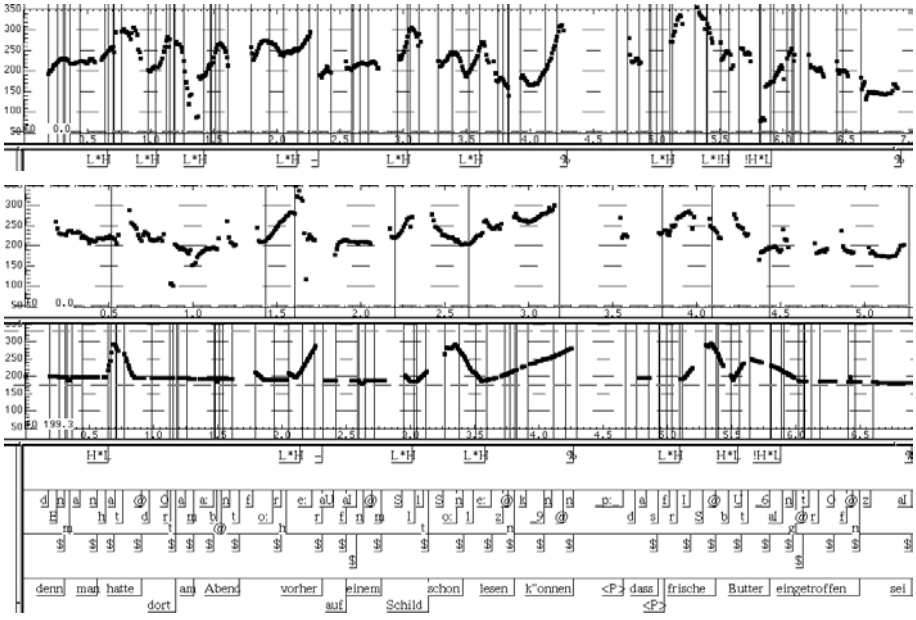

Figure 3. F_0 contour of the utterance "... denn man hatte dort am Abend vorher auf einem Schild schon lesen können, daß frische Butter eingetroffen sei" (... because one had been able to read on a sign there already the evening before that fresh butter had arrived). Top contour: American speaker's version; middle contour: German speaker's version; bottom contour: generated version of American speaker's original following the label description of the German speaker's version.

The example utterance includes a transfer of an American English continuation rise on "lesen können" which involves a rise-fall-rise movement (using the American ToBI tone description this pattern would be described as L+H* L-H%; it cannot be transcribed precisely with the German ToBI tone inventory), whereas the German version only contains a simple rise on "lesen" which spreads to the phrase boundary. The example also illustrates differences in the general placement and choice of tonal categories (e.g. a rising pitch accent on "Butter" (L+!H*) in the American version opposed to a fall (H*L) in the German version). This phenomenon extends to an overall impression of more tonal movement in the American speaker's version with additional L*H pitch accents on "dort" and "Abend". Differences in the phonetic realization of a category, e.g. in the steepness of the rises, are more subtle and can at best be hinted at in this representation of the F_0 contour.

3. Perception experiments

Any insights into the identification of actual instances of intonational foreign accent as well as into the general contribution of prosodic characteristics to the phenomenon of foreign accent must of course be validated by a substantial number of judgments from the native speakers of the respective languages. Accordingly, separate perception tests were devised for the native speakers of German and American English. In order to grant raters comfortable access to the test in terms of the flexibility when and where to take it, the perception experiment was carried out via the world wide web (go to the internet address http://www.ims.uni-stuttgart.de/phonetik/pertest/matthias.de/ for the German test or http://www.ims.uni-stuttgart.de/phonetik/pertest/matthias.ae/ for the American English test). This also had the advantage of facilitating the presentation of sound files (in the WAV format) and the processing of results. The test is self-paced and takes approxi-

mately 45 minutes to complete, but to encourage participation it was made possible to take the individual tasks separately, i.e. to leave and re-enter the test after each completed task.

This study presents the results of 57 German speakers' perception of German utterances produced by native speakers of American English.

3.1. Perceptual validation of the relevance of intonational deviations

The validation of the identified concrete instances of intonational foreign accent and consequently the described types constitutes a major aim of the perception experiments, as it also implies an examination of the general relevance of intonation to the perception of foreign accent.

The tests repeat the methodology of the analysis-by-synthesis approach in that the listeners are presented with non-native speakers' original, foreign-accented utterances and alternative versions improved by means of F_0 generation and resynthesis. The stimuli are presented in two different tasks.

In the first task the procedure of the original analysis-by-synthesis approach is simply repeated with a greater number of judges. Original and rule-generated/resynthesized versions are compared directly with listeners having to decide which rendition of an utterance exhibits less foreign accent. Thus, in a direct comparison between original and resynthesized version of an utterance, the latter is expected to be better, i.e. to exhibit less foreign accent, thereby confirming the relevance of the "corrected" deviations to the perception of foreign accent.

In the test for German listeners a significant number of the participating raters did indeed judge the resynthesized versions in all twelve pairs of stimuli to exhibit less foreign accent than the corresponding originals.

The significance of the listeners' judgments was determined by testing the identification rate of every stimulus against a binomial distribution of 0.50 (representing a chance distribution of answers, i.e., guessing), assuming a significance level p of 5%. As the listeners' answers were categorical variables on a nominal scale, consisting of only two different values (0 as a label for the original utterance - 1 for the generated version) the performance of an analysis of variance (ANOVA) to determine significance was unsuitable, as it cannot be applied to this lower-order type of scale.

In the second task listeners are given the opportunity to directly rate the degree of foreign accent. The stimuli are evaluated on a visual analogue scale, in short VAS (Wewers and Lowe, 1990) which consists in a straight vertical line with only the two extreme boundaries labelled. The scale used in this study, as shown in Fig. 4, is unipolar, denoting the absence of a phenomenon, a degree of foreign accent of 0 at one end and its maximum intensity (degree of 10) at the other end. Any click on the scale inbetween these extremes is assigned to the corresponding numerical value. There are no further marks, so-called descriptors, on the scale, which may allow the rater less possibilities of orientation, but is assumed to provide a more uniform distribution of scores along the scale's entire length (Scott and Huskisson 1976).

The VAS has the advantage of being easily grasped and requiring little motivation. They are quickly filled out and scored and can reflect fine discriminations. Apart from the already mentioned small extent of orientation they allow the rater, a possible weakness of VAS as used in this perception test lies in the indistinct definition of the scale maximum indicating the highest degree of foreign accent, allowing different interpretations by different raters. However, the absolute values for overall foreign accent are not of prime importance for this study anyway. As its purpose is rather the investigation of intonational foreign accent, it is only the comparison of those utterances that are identical in every respect except for intonation that is relevant.

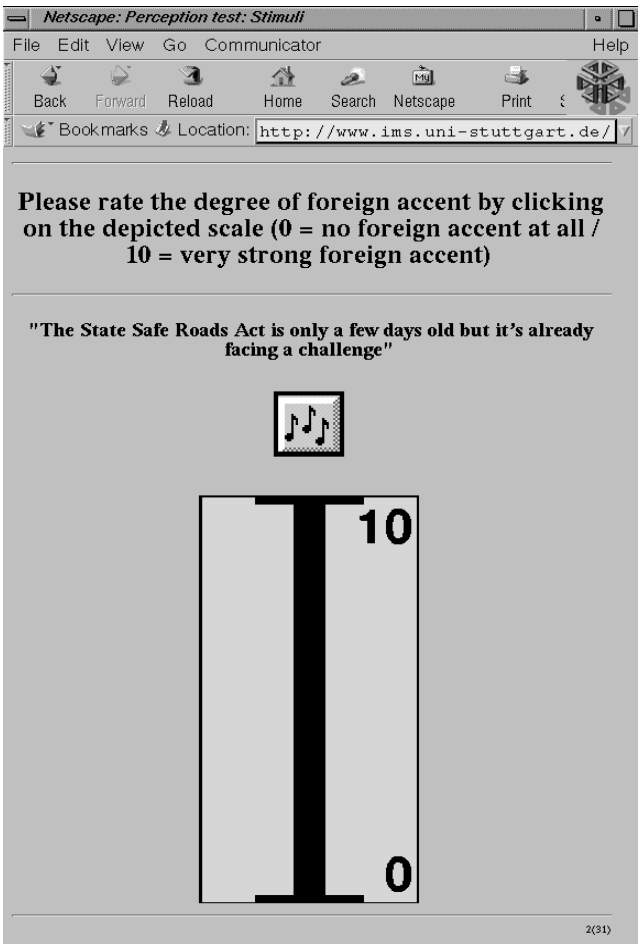

Figure 4 Rating of foreign accent on a visual analogue scale (taken from test for Americans)

The ratings were normalized using z scores. In 11 of the 15 presented pairs of original and rule-generated versions of the same utterance the generated version received a better, i.e. lower, foreign accent rating. In 9 cases the difference was significant as determined by means of paired t-tests.

Due to the random presentation of corresponding stimuli it is much more difficult for the listeners to notice and remember differences between two versions of the same utterance. It is more likely that listeners will evaluate every stimulus independently as if it were completely new without making reference to any corresponding stimuli they may have rated earlier. The ratings will thus convey the listeners' general impression of foreign accent of an utterance, somewhat neglecting smaller deviations that are only perceived in an immediate comparison of two stimuli.

For this reason it was to be expected that the difference between original and rule-generated versions would not be as obvious as in the direct pair comparison.

In summary the results of both tasks confirm the preceding analysis of intonational deviations and types of intonational foreign accent. As a consequence it is also shown that intonation does indeed make a relevant contribution to the perception of foreign accent.

3.2. The contribution of prosody in general

In order to get an impression whether prosody influences the overall perception of foreign accent, a number of original utterances of non-native speakers were low-pass filtered. Since low-pass filtered stimuli do not allow the listeners to take any segmental information into consideration, it is speculated that the prosodic properties of the presented utterances, such as fundamental frequency, stress, rhythm or speaking rate, still reflect the speakers' native languages and that therefore, for example, the English-language utterances spoken by Germans will be identified as German.

A cut-off frequency of 10% above the maximum F_0 in the phrase was used to generate the stimuli.

The test involves three different tasks: language identification, native speaker identification and pair comparison. In the language identification task the listeners are asked to decide whether the presented stimulus is spoken in German or American English. The eight stimuli used in this task are composed of four utterances with a German speaking American English, two utterances with Americans speaking German and one utterance respectively with speakers producing their own native language. The intention behind this choice of stimuli was to check whether the influence of the speakers' L1 on prosodic factors is so strong that prosody remains virtually unchanged and consequently reflects the characteristics

of the speakers' L1 only.

The significance of the speakers' classifications was again determined by testing the identification rate against a binomial distribution of 0.50.

The results show that the German listeners were generally not able to distinguish German and American English in the low-pass filtered stimuli. Only three of the eight presented stimuli yielded significant identification rates. An informal observation of the data, however, indicated a possible correlation between stimulus duration and language identification rate and indeed Pearson's correlation coefficient showed a significant ($p = 0.041$) correlation of 0.767 between the logarithm of stimulus duration and the language identification rate. To further confirm this connection an additional test was devised which contained only low-pass filtered stimuli between 35 s and one minute in duration. The additional test was taken by 34 German listeners with identification rates now being significant for five of seven presented stimuli.
In the native speaker identification task listeners are asked to decide whether a presented low-pass filtered stimulus in their L1 (text given) was produced by a fellow native speaker or a non-native speaker (American or German respectively). The 57 German listeners were able to correctly and significantly decide whether they heard a native speaker or not in six of the eight presented stimuli.

In the final task the raters were given the opportunity to directly compare two versions of the same utterance as produced by a native speaker and a non-native speaker. It was expected that prosodic differences would be particularly obvious under these conditions, and indeed a significant number of the participating listeners was able to distinguish the German speaker from his/her American counterpart in all presented stimuli with the exception of a control stimulus which opposed the productions of two German speakers.

The results of these tasks can be interpreted as showing that listeners have clear ideas as to which prosodic characteristics are possible and appropriate in their native language and which are not or at least less so. Therefore it can be stated that prosody contributes considerably to the perception of foreign accent.

3.3. The contribution of intonation in comparison with that of other prosodic aspects

While the previous tasks have aimed to demonstrate the relevance of prosodic, in other words non-segmental, aspects to foreign accent, they do not reveal if, and if so to what degree, prosodic characteristics other than intonation contribute. For this reason, in an additional test, original signals are first given a continuous monotonous intonation (e.g., 220 Hz for female voices) and then low-pass filtered. This kind of signal includes all prosodic properties, especially rhythm and speaking rate except everything associated with fundamental frequency. The idea is that the influence of intonation can be determined by a comparison of the regular low-pass filtered version of a stimulus and the monotonous low-pass filtered version. The different ratings for the otherwise identical stimuli should allow an evaluation of the effect of intonation on the perception of foreign accent in relation to that of the other prosodic features.

The perception test results for these stimuli show that in the direct pair comparison stimuli that lack intonational information but retain all other aspects of prosody are still sufficient for a significant number of listeners to recognize the version produced by the native speaker of their language. However, these results are nevertheless significantly worse than those yielded by the regular low-pass filtered stimuli. In addition to this, the low-pass filtered stimuli with monotonous intonation also show no significant results for the language identification task, no matter if for regular or long stimuli, or for the native speaker identification task. Indeed, the recognition rates are again significantly worse than they are for the same tasks using low-pass filtered stimuli with intonational information.

The significance of these differences is determined by testing the recognition rates of the monotonous low-pass filtered stimuli against a binomial distribution that is identical to the recognition rate of the corresponding "regular" low-pass filtered stimuli.

In summary, it can thus be stated that intonation constitutes the crucial, significant part of the totality of prosodic functions.

3.4 Comparing the contribution of segmental and intonational deviations to the perception of fo-reign accent

The conducted perception tests do show that intonation is relevant to the perception of foreign accent, but not how important it is in relation to segmental foreign accent, i.e., how big its contribution to an overall effect of foreign accent actually is.

There is a widespread view that segmental foreign accent is perceived much more strongly, leading to it very often being equated with foreign accent per se, especially in a comparison of stress accent languages such as German and American English.

A first attempt to describe the relative importance of tonal and segmental deviations is undertaken in this study, as the foreign accent ratings for two different versions of the same stimulus are compared. The crucial difference between the two versions consists in the fact that one is meant to show segmental, but not intonational foreign accent (F_0 generated and resynthesized version), while the other has intonational, but not segmental foreign accent (fully synthesized version using diphones from the MBROLA project (Dutoit et al., 1996)). The differences in the ratings were evaluated by means of paired t-tests, showing that the synthesized stimuli without segmental foreign accent did indeed receive better ratings, the difference being significant in two of the four compared pairs.

While this type of comparison does not directly quantify the share of segmental and intonational foreign accent in

the perception of foreign accent in general, it nevertheless confirms that the contribution of segmental deviations is much more important. This is emphasized by the fact that the synthesized stimuli used in the experiment do not sound completely natural, which is likely to have a negative influence on the ratings. Also, it must not be forgotten that only stimuli were evaluated that actually did exhibit intonational foreign accent. As mentioned before, this is not necessarily the case in many of the other productions of non-native speakers. On the other hand it could be argued that the stimuli in the foreign accent rating task do not take into account the effects of accumulation of slight tonal or rhythmic deviations in combination with the overall discourse context. These factors might increase the contribution of prosodic characteristics to the impression of foreign accent in natural, everyday discourse.

For this reason it has to be stressed again that this particular test can only be considered a first attempt to describe the relationship between intonational and segmental aspects of foreign accent.

4. Conclusion

With regard to the overall aim of this study, which was to provide a perceptual analysis of the contribution of prosody in general and intonation in particular to the perception of foreign accent, the results of the performed perception experiments can be summarized as follows:

- prosody significantly contributes to an overall impression of foreign accent
- intonational aspects are relevant to the perception of foreign accent
- intonation is the most influential prosodic aspect in foreign accent
- segmental aspects of foreign accent are perceived more strongly than intonational ones

Additionally, the study is based on the successful use of speech technological applications such as F_0 generation and resynthesis to identify concrete cases and categories of intonational foreign accent which are then validated in the perception tests. Important basic assumptions as to the behavior of intonational deviations are also formulated.
The study thus provides a first overall description of the roles prosody and especially intonation play with respect to the phenomenon of foreign accent in German and American English.

5. References

Beckman, M. and G. Ayers (1994): *Guidelines to ToBI Labelling*. Version 2.0. Ohio State University

Best, C. T. (1995): A direct realist view of cross-language speech perception. In: W. Strange, (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 171 - 204).Timonium, MD: York Press

Dutoit, T., V. Pagel, N. Pierret, F. Bataille and O. van der Vreken (1996): The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes. In: *Proceedings of the International Conference on Spoken Language Processing* (pp. 1393-1396*)*. Philadelphia

Flege, J. E. (1995): Second language speech learning: theory, findings and problems. In: W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 233-277). Timonium, MD: York Press

Grabe, E. (1997). Comparative intonational phonology: English and German. In: *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications* (pp. 157 - 160). Athens

Grabe, E. (1998): Pitch accent realization in English and German. In: *Journal of Phonetics 26* (pp. 129 - 143)

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Doctoral Dissertation, Arbeiten des Instituts für Maschinelle Sprachverarbeitung (AIMS) Vol. 6(3), University of Stuttgart

Jilka, M., G. Möhler and G. Dogil (1999): Rules for the generation of ToBI-based American English intonation. In: *Speech Communication 28* (pp. 83 - 108)

Kohler, K.J. (1994): *The Kiel corpus of read speech, Vol. 1*. CD-ROM, Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel

Markham, D. (1997): *Phonetic imitation, accent, and the learner*. Doctoral Dissertation. Travaux de l'Institut de Linguistique de Lund 33, Lund University Press

Ostendorf, M,, P. J. Price and S. Shattuck-Hufnagel (1995): *The Boston University Radio News Corpus*. Technical Report ECS-95-001, Electrical, Computer and Systems Engineering Department, Boston University, Boston, MA

Pierrehumbert, J. (1980): *The Phonology and Phonetics of English Intonation*. Ph.D. Dissertation, M.I.T.,Cambridge, MA

Rapp, S. (1998): *Automatisierte Erstellung von Korpora für die Prosodieforschung*. Doctoral Dissertation, Arbeiten des Instituts für Maschinelle Sprachverarbeitung (AIMS) Vol. 4(1), University of Stuttgart

Scott, J. and E.C. Huskisson (1976): Graphic representation of pain. In: *Pain 2* (pp. 175-184)

Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert and J. Hirschberg (1992): ToBI: A standard for labelling English prosody. In: *Proceedings of the 1992 International Conference on Spoken Language Processing* (pp. 867 - 870)

Wahlster, W. (1997): VERBMOBIL: *Erkennung, Analyse, Transfer, Generierung und Synthese von Spontansprache*. Report-198-9.DFKI GmbH

Wewers, M. E. and N.K. Lowe (1990): A critical review of visual analogue scales in the measurement of clinical phenomena. In: *Research in Nursing and Health 13* (pp. 227-236)