INTONATIONAL VARIATION AND PRONUNCIATION APTITUDE

Von der Philosophisch-Historischen Fakultät der Universität Stuttgart

zur Erlangung der Würde eines Doktors der

Philosophie (Dr. phil.) genehmigte Abhandlung

Vorgelegt von

Volha Anufryk

aus Minsk

Hauptberichter: Prof. Dr. phil. habil. Grzegorz Dogil

Mitberichter: Prof. Dr. phil. habil. Bernd Möbius

Tag der mündlichen Prüfung: 02.12.2011

Institut für maschinelle Sprachverarbeitung

Universität Stuttgart

2011

Erklärung

Hiermit erkläre ich, dass ich, unter Verwendung der im Literaturverzeichnis aufgeführten Quellen und unter fachlicher Betreuung, diese Dissertation selbständig verfasst habe.

(Volha Anufryk)

ACKNOWLEDGEMENTS

In Russian we say that you have to be lucky in two things: with your parents and with your teachers. With this criterion, I think I can consider myself a very lucky person.

My mother has supported me as long as I can remember, sharing any effort, burden, just sharing.

Apart from her, I have had amazing teachers. I will always be thankful to Professor Grzegorz Dogil for running the risk in taking me on when he knew nothing about me. He has continued to amaze me over the years of our acquaintance with his ability to find inspiring ideas in virtually any area, even seemingly the ones he is not particularly familiar with.

In my earlier years, since the first year as an undergraduate student, I worked very closely with Professor Elena Karnevskaya, whom I owe the basis of my theoretical knowledge and critical thinking in Phonetics.

I have also learned a lot from my colleagues and professors of the Institute for Natural Speech Processing during my time as a PhD student. The work of my second advisor, Professor Bernd Möbius, has lent me valuable ideas for this dissertation, and his advice has helped to move forward with it.

A creative period is related to my participation in the “Language Talent and Brain Activity” project with Dr. Matthias Jilka, Dr. Susanne Reiterer, Dr. Giuseppina Rota, Henrike Baumotte, Natalie Lewandowska and Dr. Jagoda Sieczkowska. Not only has this project inspired the main idea of this dissertation, it was always great fun working and spending time with these people.

I am also deeply thankful to the Graduate College of Universität Stuttgart "Sprachliche Repräsentationen und ihre Interpretation" and personally Professor Jürgen Pafel for the support they have given me for my research. Through the Graduate College I have also met a lot of interesting people, fellow PhD students, with whom we not only shared ideas during our yearly meetings in Kleinwalsertaal, we also became friends. This way I met my good friend, now Doctor, Katja Lapshinova-Koltunski.

Finally, I am grateful to my boyfriend Michael Weaver for putting up with me during this time, as well as really helping with some programming for statistical analysis.

All these people and many others have helped me achieve one of the greatest goals of my life. And for that I would like to say – thank you.

TABLE OF CONTENTS

Table of contents

SUMMARY

The current dissertation is an investigation of intonational variation. It was inspired by the core idea that variability of prosody is related to pronunciation aptitude. We based this idea on the analogy with other linguistic domains: speakers are considered to be talented and effective, in case they possess and employ a great variety of syntactic structures and vocabulary items. Following this logic, we hypothesized that the more talented a speaker / learner is, the more prosodic variability should be observed in his or her speech, both in the native language and an L2. A review of the major general linguistic and second language acquisition theories presented in the given dissertation appeared to support this general assumption.

We therefore conducted an experiment, in which we investigated productions of 41 native German speakers varying in pronunciation aptitude – low-aptitude, below-average, average, above-average, high-aptitude and highest-aptitude. This aptitude classification originated from an earlier study by Jilka (Jilka 2009; Jilka et al. 2007). The subjects performed read and spontaneous speech tasks in English and German, and their realizations were further compared with the samples produced by a pool of 12 native English speakers, as to the degrees of variability.

With the current dissertation, we therefore present an interdisciplinary study, linguistic and SLA, comprising three major underlying aspects:

- *theoretical*, discussed in Chapters 1-2, laying down general linguistic and SLA conceptual grounds for the core hypotheses put forth in Chapter 3 and for the subsequent experimental endeavours;
- *methodological***,** presented in Chapters 4-5, explaining the choice of the instrumental paradigm and the general methodology;
- *experimental*, described in Chapters 6-7, dwelling upon the results of the above-mentioned study of intonational variability in the realizations of 41 native German speakers; an experiment based on the theoretical and methodological premises established in the earlier dissertation chapters.

The three fundamental aspects were further confronted with one another in Chapter 8, i.e. the theoretical and methodological assumptions were brought in line with the results of our experiment.

Further, we will briefly discuss the major research aspects, as well as some additional issues raised in the current dissertation.

The methodological aspect

A thorough examination of intonation models (theoretical, experimental-phonological and experimental-phonetic) in Chapter 4 allowed us to select the general methodological framework of autosegmental-metrical phonology and its instrumental paradigm, the ToBI labeling convention, as the basis for the phonological analysis in the current study. This choice was explained by the ability of the model in question to account for phonological variability of intonation, viewed as sequences of single intonation events (e.g. Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Hirschberg & Pierrehumbert 1986; Gussenhoven 1983; Ladd 1998), whose exact inventories for a specific language are stipulated in the ToBI.

Bearing in mind that ToBI conventions exist for both the German and English languages (Silverman et al. 1992, Pitrelli et al. 1994; Grice & Benzmueller 1994; Mayer 1995), this paradigm seemed suitable for our research purposes. However, due to the fact that we analyzed German-accented speech in English, where it was not always possible to assign a given realization to an English ToBI category, we opted for a combined inventory of both English and German ToBI labels. The corresponding labeling approach is described in Chapter 5.

Autosegmental-metrical phonology and ToBI allow of a phonological description of intonational variation. However, an all-round approach to the latter phenomenon presupposes the investigation of the phonetic level as well. We therefore reviewed the major existing phonetic models of intonation (Öhman's superpositional model **-** Öhman 1967; Fujisaki's model **-** Fujisaki 1983; Van Santen and Möbius' quantitative model of *F0* generation and alignment - Van Santen and Möbius 2000; Tilt Model - Taylor 2000) and concluded that the PaIntE paradigm (Möhler 1998, 2001) most closely approaches the instrumental requirements of the current study. Its principal advantage over other *F0* approximation models is the fact that only the fine phonetic detail, induced by a given segment (phoneme) is removed from the contour, making it even more perceptually acceptable than the original one, according to Möhler's perceptual experiments. Apart from this advantage, it is also possible to apply the PaIntE model to individual ToBI labels, which creates a phonetic / phonological compatibility for analysis.

PaIntE describes an individual *F0* curve in terms of six basic parameters:

- ***d*** – the amplitude of the *F0* peak in Hertz;
- ***b*** – the alignment of the *F0* peak – a relative measure between -1 and 2, indicating the location of the peak within a three-syllable window;
- ***a1*** and ***a2*** – the relative velocity of the rising and falling sigmoids of an *F0* curve, respectively;
- ***c1*** and ***c2*** – amplitudes of the rising and falling sigmoids of an *F0* curve, respectively, in Hertz.

In the course of data analysis we analyzed intonation in terms of *F0* curves carrying a phonological ToBI label, or labels, with the phonetic properties stipulated by the PaIntE parameters.

The theoretical aspect

From the theoretical perspective, we explored the major SLA and general linguistic theories, in order to prove or disprove our initial assumption of a possible correlation between intonational variation and pronunciation aptitude. We also strived to find out, which theories best account for the two phenomena in question.

For the *general linguistic theories*, a conclusion was drawn that Exemplar Theory (e.g. Nosofsky 1986, Pierrehumbert 2001, Bybee 2005, Wade et al. 2010) can predict intonational variation most accurately, due to the fact that it takes into account concrete speech representations, allowing of realistic descriptions of intonational categories. The experimental study described in Chapters 6-7 did not appear to contradict with this initial assumption. However, the outcome of the current investigation implied that Exemplar Theory is perhaps more suitable for the description of infrequent categories, rather than frequent ones, due to the instability and unpredictability of the former type of phonological events.

Thus, the simple low L* pitch accents as well as intermediate low L- and high H- boundary tones, registered as infrequent in the course of the current investigation, did not suggest any clear distribution patterns, either on the phonetic or the phonological level.

When analyzing the performance of the six aptitude groups, we also noticed that the larger group samples (frequent categories) exhibited relatively stable variability and frequency of occurrence patterns throughout the course of the study, whereas one-person groups (infrequent categories) often deviated from the overall trends in an inconsistent way. Thus, the number of tokens produced by the above-average female and the below-average male informants, one-person groups, was smaller than in the remaining groups, predicting a lower degree of variability. However, one-person group samples exhibited both lower and higher, as well as equal, degrees of phonetic variability as contrasted to the other aptitude groups.

Analyzing variability effects in an infrequent category, one might notice that in cases with high degrees of variability, individual tokens may be located at a considerable distance from one another, coinciding with or even exceeding the range of values in a frequent category, with a denser population of tokens. It appears that the investigation of infrequent categories in terms of individual exemplars could shed more light into the functioning of these language events, since generalization in terms of range is rather problematic.

More instances of idiosyncratic behaviour within the single-person groups were found on the phonological level, especially with respect to the below-average male test person. He employed the rising tone category L*H and the high full boundary tone category H-H% least extensively of all female and male L1 German groups. On the contrary, the percentage of H* targets in the below-average male speaker's realizations exceeded that in the remaining groups.

Evidently, the above examples point out to the instability of single-person groups (infrequent human categories), which may, on the one hand, comply with the more general sample or depart from it, on the other. This fact, in turn, is also related to the notion of individual variation and the heterogeneity of human populations, where some individuals would be perfect representatives of general trends, while others would render outlying instances. Further on, such outliers may, with category evolution, become regular category members. It appears in this regard that Exemplar Theory would be most suitable for the description of infrequent categories and their development, as it is meant to analyze individual speech representations.

For frequent category mechanisms, Structural Phonology (e.g. Saussure 1916; Trubetzkoy 1958; Bloomfield 1933) provides some valid explanations. Once a frequent category is established, in all its aspects, its members populate a general common space. Inside this space, according to the structural phonological approach, one can establish boundaries, within which exemplars can be freely produced. However, any exemplars emerging beyond the category boundaries during its development would not be categorized. The structural approach thereby appears more suitable for synchronic descriptions of frequent categories, while Exemplar Theory takes account of the diachronic aspect as well.

The *SLA theories* reviewed in Section 2.2 also found some support of their underlying ideas in the current investigation.

We could therefore relate some of the experimental results to the *Basic Variety Theory* (Klein & Perdue 1997). According to BV, SLA is characterized by a gradual increase in the degree of speech variability as the learner's proficiency reaches higher levels, gradually approaching the target-language variation levels, i.e. native-speaker performance.

The current investigation but partially confirms this idea on the phonetic level. Indeed, the degree of phonetic variability demonstrated by the highest-aptitude and high-aptitude female groups was greater as compared to the other L1 German female speakers. At the same time, such a high degree of variation was also closest to the native female group sample, which was evident in several experimental conditions (e.g. *c1* parameter values of L*H accents analyzed by speech type and gender; parameter *d* analysis of H* events by gender and in female read speech in both languages – see Chapters 7-8 for a full overview).

Such proximity of higher-aptitude and native speakers was perhaps the sole consistent evidence in favour of the Theory of Basic Variety. The other results led to contrary conclusions. Thus, although we could group female subjects with respect to the degrees of variability, this pattern was not supported by male speech.

Another counterargument to the BV is the absence of any gradual increase in the degrees of variability with the rising level of pronunciation aptitude. Within the female subjects, to whom the theory in question is only applicable, only the highest-aptitude and high-aptitude groups were characterized by consistently more variable samples.

Selinker's *Interlanguage Theory* (Selinker 1972) and its major concepts provide a relatively solid explanation to SLA processes. One of the most pertinent concepts for the given study is the effect of transfer. We could indeed register instances of consistent phonological transfer from L1 into L2, which was most characteristic of the events ending on high pitch. Rising tones L*H and the high full boundary tones H-H% enjoyed a greater frequency of occurrence, caused by L1 transfer, within native German speakers. This appeared to be one of the cardinal tendencies of the given investigation due to the greater distribution of the events at issue in German (e.g. Mayer 1995) and their much sparser employment in English.

This tendency was, however, suppressed rather consistently by female subjects of highest and high pronunciation aptitude. In their L2 productions, these speakers reduced the number of rising L*H accents and H-H% boundary tones. This, in turn, appeared to approach the two higher-aptitude female groups to the native speakers. Male productions did not support the tendency in question in a consistent manner. Summarizing, we observed a transfer of a typical category by most female groups into L2, resistance to the transfer phenomenon within the highest-aptitude and high-aptitude female speakers as well as their accommodation to the L2 frequency of occurrence pattern.

The *Diffusion Model* (Gatbonton 1978) postulated that newly acquired L2 forms coexist with L1 exemplars in free variation, and later they are sorted into their appropriate contexts. We could not disprove this general postulation with the experimental results, as far as the phonetic level of analysis is concerned. As mentioned previously in the interpretation of the Theory of Basic Variety, we could only separate the highest-aptitude, high-aptitude and native female subjects with consistently higher degrees of variability.

Some explanation of the SLA processes can be traced back to the *Perceptual Assimilation Model* (Best 1995). PAM posits that in typologically close languages a counterpart / close L2 category tends to merge with the corresponding L1 event. In English and German, often been reported as typologically close languages with respect to intonation (e.g. Kuhlmann 1952, Schubiger 1965, Esser 1978, Scuffil 1982, Grabe 2000), L*H and H-H% categories might have been assimilated by most groups across L1 and L2, due to their active consistent transfer into English. A perceptual investigation is needed, however, in order to corroborate such assimilation.

The *Second Language Speech Learning Theory* (Flege 1995) also sheds some light onto the SLA processes at play in the current investigation. It does, therefore, explain the reason for L1 transfer into L2. In Flege's opinion, typological proximity hampers category discrimination. Consequently, the undiscriminated L1 / L2 categories undergo transfer, as did, for instance, the rising tones L*H and high full boundary tones H-H% in the realizations of most aptitude groups in the current study.

In general, we can state that most SLA theories account for the mechanisms guiding the SLA processes to a greater or lesser extent. However, the Interlanguage Theory appears to be the closest predictor of cross-language intonational variation.

The experimental aspect

One of the most crucial parts of the current investigation consisted in the validation, or rejection, of the hypotheses presented in Chapter 3. All of these hypotheses originated from the underlying idea of a correlation between pronunciation aptitude and intonational variation. In our study, the hypotheses applied more specifically to the investigated aptitude / proficiency groups. We will briefly recount each of the hypotheses and discuss the experimental outcomes pertaining to them.

The first hypothesis stated that L2 speech of less proficient / less talented speakers should exhibit a smaller degree of variability in a foreign language. In the experimental data we could see that subjects of highest and high pronunciation aptitude, especially female, consistently demonstrated higher degrees of variability than the remaining groups (e.g. in *c1* parameter values of L*H accents analyzed by speech type and gender – see Chapters 7-8 for a full overview). Therefore, we might consider the first hypothesis, for the most part, validated.

In *the second hypothesis* we posited that high-aptitude / high-proficiency individuals should be more variable in their speech due to their higher acuity in mastering L2 exemplars and employing them in communication. Like the first hypothesis, evidence for the current assumption was provided by the performance of the two higher-aptitude female groups in their realization of phonetic PaIntE parameters in a number of experimental conditions (see Chapters 7-8 for a full overview).

Logically following from the second hypothesis, we further assumed in *hypothesis three* that average-proficiency / average-aptitude speakers are expected to exhibit average variability levels, relative to the higher degrees of variation within the more talented speakers and less variable samples of the less talented individuals. This idea was partially rejected by the experimental results on the phonetic level: the average subjects, especially female, exhibited variability degrees close to those of the highest-aptitude and high-aptitude groups in several conditions (e.g. parameter *d* of L*H accents: by speech type in both language conditions, in the speech of average informants in the whole experimental corpus in L1 and L2 and in English female read speech – see Chapters 7-8 for a full overview).

The *fourth hypothesis* was also related to average-aptitude / average-proficiency speakers. We thereby assumed that degrees of variation between speakers of average and higher than average pronunciation aptitude are unpredictable and can only be verified through the analysis of real-life speech data. As seen in the corpus, only the highest-aptitude and high-aptitude female subjects were singled out for their greater degrees of variability in the realization of PaIntE parameters. The variability ratio between the remaining groups was not distinct and their performance was not consistent. Therefore, the given hypothesis proved true for most aptitude groups, and not only those of average and higher than average aptitude / proficiency.

The final hypothesis, *hypothesis five*, postulated that on a proximity scale, speakers of higher than average aptitude and proficiency should approach the native-speakers' variation pattern most closely. Average as well as lower than average levels of aptitude and proficiency should correlate with lower degrees of variability, establishing a greater distance of such speakers from the proper L2 performance. Supporting the first part of the hypothesis, highest-aptitude and high-aptitude female informants displayed higher degrees of phonetic variability than the other L1 German female groups, which was also closest to the native subjects in performance on the phonetic level in a number of conditions (e.g. parameter *d* analysis of H* events by gender and in female read speech in both languages – see Chapters 7-8 for a full overview).

To summarize the results of the hypotheses validation, we ought to point out that the experimental evidence presented above only fully discarded the second hypothesis, while supporting the remaining ones. However, these results only relate to the phonetic level. The phonological analysis, on the other hand, did not appear to follow these assumptions. We will therefore further discuss the implications of the phonological investigation, alongside with any additional phonetic issues.

Additional experimental phonetic and phonological aspects

The *analysis of the ToBI pitch accents and boundary tones* allowed us to relate the degrees of variability within the aptitude groups to the L1 / L2 accommodation vs. transfer phenomena. As a result, we found that most L1 German groups tended to transfer their L1 behaviour to English, whereas speakers of highest and high pronunciation aptitude attempted to accommodate to the L2 model. This tendency was especially prominent in the L*H pitch accent and H-H% boundary tone categories, where the highest-aptitude and high-aptitude informants reduced the percentages of these typical German categories in their L2 speech, as compared to German, where these events are more widely distributed. This made their performance closest to the native subjects, also exhibiting even lower numbers of rising tones and high full boundary tones.

As evidenced from the experimental results, the phonetic and phonological levels follow a different logic with regard to the degrees of variability in individual pitch accents and boundary tones. The phonetic results relate to the quantitative aspect of variability – its greater / smaller degrees; the phonological analysis rather reflects the interlanguage transfer and accommodation phenomena in intonation. The analysis of the corpus on the text level appeared to suggest still another pattern across the aptitude groups.

During the *global phonological investigation on text level*, we compared ToBI transcriptions of the two texts under analysis, the fable "The North Wind and the Sun" and the abstract from "Mrs McWilliams" by Mark Twain, for the uniformity / variability of prosodic text interpretation. Thus, the first tendency was the lower degree of cross-speaker variability in all German-language transcriptions of L1 German speakers as compared to greater variability in all their English productions. One of the logical explanations to this fact could be the general comfort of the subjects with their mother tongue and their subsequent neutral interpretation of the experimental texts, which also suggest some common text interpretation patterns in a given community of speakers

The English versions of the two investigated texts allowed us to discover some other consistent patterns of prosodic text interpretation across the groups. Thus, the native and average subjects demonstrated the highest degrees of cross-speaker agreement, regardless of the gender factor in the "North Wind and the Sun" samples as well as in female speech of "Mrs. McWilliams". One might possibly explain such results by these subjects' relatively neutral attitude to the task, which yielded neutral realization templates. The highest-aptitude and low-aptitude speakers, on the other hand, demonstrated the greatest degrees of variability in prosodic text interpretation, which was especially evident in their productions of "Mrs. McWilliams". We would assume different explanations for this fact for the two groups. The highest-aptitude informants might have strived for a more varied and "interesting" interpretation of the experimental texts, which entailed using varying intonational means and resulted in a higher degree of cross-speaker disagreement. The low-aptitude speakers, on the other hand, regarding their lower proficiency level, might not have been certain, which event to employ in this or that context, hence the higher Levenshtein distance values attesting to greater disaccord within the group in this respect.

The underlying reasons for greater / smaller degrees of cross-speaker variability in prosodic text interpretation across the groups did not appear to be guided by only phonetic or phonological factors per se. Rather, the results reflect the speaker's extra-linguistic predisposition towards the experimental task (neutral – for average and native subjects; more involved – for the highest-aptitude group) as well as their proficiency level in English (low-aptitude subjects perhaps did not possess a satisfactory inventory of intonation units). Thus, the inventory of language units at a subject's disposal, coupled with his or her knowledge of their application in an appropriate context, defines the degrees of variability. Personal predisposition appears to be a more complex subject, although positive predisposition should intuitively contribute to variability. Lower-aptitude subjects might therefore apply prosodic events inappropriately and differently from speaker to speaker, which creates a greater degree of cross-speaker variability.

An additional *phonetic* tendency to the ones discussed with respect to the experimental hypotheses as well as the general linguistic and SLA theories was the fact that female speakers of highest and high pronunciation aptitude realized their PaIntE parameters on a higher mean *F0* level than the other groups in a number of experimental conditions (see Subsection 7.2.1 for a summary). This finding could be related to the above tendency of the highest-aptitude group for a higher degree of variability in prosodic text interpretation. It appears that the subjects not only attempted to make their speech more variable phonologically, they also raised their mean pitch higher than the native speakers. Several earlier studies also referred to such a phenomenon of overhitting the L2 target (e.g. Bohn & Flege, 1992).

To conclude the summary of the experimental results, it is worthy of mentioning that some of the tendencies mentioned above concur with our previous research on the same groups of subjects. Thus, we previously encountered a higher degree of variability as well as higher pitch within high-aptitude and highest-aptitude speakers, in particular, female (Anufryk 2009; Anufryk & Dogil 2009). The tendency of these groups to accommodate to the English distribution pattern by suppressing the number of rising pitch accents and high boundary tones was also supported by our earlier investigations (Anufryk 2008; Anufryk et al. 2008; Anufryk 2009; Anufryk & Dogil 2009). Finally, for the intonational text interpretation, our previous results were also similar: less variability in all German-language samples of native German speakers; in English, more agreement between native speakers and the average group and more disagreement between high-aptitude and highest-aptitude as well as low-aptitude speakers (Anufryk 2009).

Conclusion

From the linguistic point of view, the focus of our attention lay within the frequent phonological categories, which appeared to accord mostly with the concepts of the Exemplar Theory and Structural Phonology. We concluded that the former is best-suited for a dynamic description of category functioning and formation, as it takes account of individual phonetic representations. The structural phonological approach is capable of category definition statically: any outlying instances cannot be explained.

Infrequent phonological events remained beyond the scope of the study. It appears, however, that infrequent classes provide insight into the diachronic perspective, i.e. category development, alongside with rendering synchronic information on how they function at a given point in time. It would therefore be interesting to dedicate attention to infrequent categories in our future work.

From the SLA perspective, the results of this experimental investigation fell within the predictions of several major second language acquisition theories – Basic Variety, Interlanguage, Diffusion Model, Perceptual Assimilation Model and Second Language Speech Learning Theory. The current study, however, could contribute to the understanding of their underlying principles by outlining their limitations as to some aspects – e.g. the absence of gradual increase in intonational variability with rising levels of pronunciation aptitude / proficiency, contrary to the predictions of the Basic Variety Theory.

The methodology of the current study allowed us to obtain detailed linguistic information on individual intonation events, both phonological, reflected through the ToBI labels, and phonetic – by means of the six PaIntE parameters. Apart from that, we could investigate the extralinguistic factor of pronunciation aptitude, through its manifestation in proficiency.

When we, further on, carried out an experiment with this methodology and based on the theoretical premises of the major SLA and general linguistic theories, we could find support for most of the experimental hypotheses, alongside with being able to point out to additional experimental and theoretical issues.

Chapter 1 introduces the notion of pronunciation aptitude, thus explaining the motivation behind the current investigation. Chapter 2 discusses the phenomenon of intonational variation, as it is reflected in the major general linguistic and second language acquisition theories. In Chapter 3 we present the experimental hypotheses following the implications of the linguistic theories introduced in Chapter 2. A theoretical and methodological framework for the study is selected in Chapter 4, after a discussion of a number of prosodic models. Chapter 5 dwells upon the procedure of the experiment, and Chapters 6-7 present its results. The final chapter, Chapter 8, compares the results of our investigation with the theoretical premises of the major linguistic theories. Apart from that, a validation / discarding of the experimental hypotheses is put forth. Finally, we discuss whether the results of the current study correspond with the outcomes of our earlier investigations of the phenomenon of intonational vatiation in second language speech.

ZUSAMMENFASSUNG

Die aktuelle Dissertation ist eine Untersuchung der intonatorischen Variation. Der Arbeit liegt der Kerngedanke zugrunde, dass die Variabilität der Prosodie mit dem Aussprachetalent zusammenhängt. Diese Idee stammt aus der Analogie mit anderen Sprachbereichen: man betrachtet einen Sprecher als talentiert und effektiv, falls er eine Vielzahl von syntaktischen Strukturen und Vokabeln besitzt und gebraucht. Dieser Logik folgend, schlagen wir volgende Hypothese vor: je talentierter ein Sprecher / Lerner ist, desto mehr prosodische Variabilität in seiner Rede zu beobachten ist, sowohl in der Muttersprache als auch in einer L2. Ein Überblick über die wichtigsten sprachwissenschaftlichen und SLA-Theorien in der gegebenen Dissertation schien diese allgemeine Annahme zu unterstützen.

Deswegen führten wir ein Experiment aus, in dem Produktionen von 41 deutschen Muttersprachlern untersucht wurden. Die Sprecher unterschieden sich in der Aussprachebegabung – untalentiert, unterdurchschnittlich, durchschnittlich, überdurchschnittlich, talentiert und sehr talentiert. Diese Klassifizierung entstand aus einer früheren Studie von Jilka (Jilka et al 2007; Jilka 2009). Die Probanden hatten Aufgaben in spontaner und gelesener Sprache auf Deutsch und Englisch ausgeührt, und ihre Produktionen wurden weiter mit den Proben eines Pools von 12 englischen Muttersprachlern im Grad der Variabilität verglichen.

Die aktuelle Arbeit ist daher eine interdisziplinäre Studie, linguistische und SLA, die aus drei großen zugrundeliegenden Aspekten besteht:

- *Der theoretische Aspekt* wurde in Kapiteln 1-2 diskutiert. Die allgemeinen linguistischen und SLA-Prinzipien wurden festlegt, die zur weiteren Formulierung der Kern-Hypothesen im Kapitel 3 dienten;
- *Der methodische Aspekt* wurde man in Kapiteln 4-5 vorgestellt, dabei wurde auch die Wahl des instrumentalen Paradigma und der allgemeinen Methodologie erklärt;
- *Der experimentelle Aspekt* basierte auf den Ergebnissen der oben erwähnten Studie der intonatorischen Variation in den Realisierungen von 41 deutschen Muttersprachlern.

Die drei grundlegenden Aspekte wurden danach miteinander im Kapitel 8 konfrontiert, d.h. die theoretischen und methodischen Annahmen wurden im aktuellen Experiment getestet.

Weiterhin werden wir kurz die Hauptaspekte der Forschung sowie einige weitere in der aktuellen Dissertation erhobene Probleme zusammenfassen.

Der methodische Aspekt

Eine gründliche Untersuchung der Intonationsmodelle (theoretisch, experimentell-phonologisch und experimentell-phonetisch) im Kapitel 4 hat es uns ermöglicht, die allgemeinen methodischen Rahmen der autosegmentalen-metrischen Phonologie und ihres instrumentalen Paradigmas, der ToBI Konvention, als Grundlage für die phonologische Analyse in der aktuellen Studie auszuwählen. Diese Wahl kann man durch die Fähigkeit des Modells erklären, die phonologische Variabilität der Intonation als Sequenzen einzelner Intonationsereignisse zu betrachten (z.B. Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Hirschberg & Pierrehumbert 1986; Gussenhoven 1983 Ladd 1998). Die genauen Inventare der Intonationseinheiten für eine bestimmte Sprache sind in der ToBI Konvention festgelegt. Wichtig für die aktuelle Studie ist die Tatsache, dass ToBI Systeme für Deutsch und Englisch (Silverman et al 1992, Pitrelli et al. 1994; Grice & Benzmüller 1994; Mayer 1995) existieren, deswegen schien dieses Paradigma als adäquat für unsere Forschungszwecke zu sein. Doch da wir die englische Sprache mit einem deutschen Akzent untersuchten, war die Kategorisierung zu einer englischen ToBI-Kategorie nicht immer möglich. Deswegen entschieden wir uns für ein kombiniertes Inventar deutscher und englischer ToBI-Labels, das im Kapitel 5 beschrieben ist.

Die autosegmentale-metrische Phonologie und ToBI bieten eine phonologische Beschreibung der intonatorischen Variation an. Allerdings ist die Untersuchung der phonetischen Ebene genauso wichtig für eine ausgewogene Analyse. Deshalb haben wir die wichtigsten existierenden phonetischen Modelle der Intonation (das Fujisaki Modell - Fujisaki 1983; das Van Santen und Möbius quantitative Modell der *F0*-Generation und Ausrichtung - Van Santen und Möbius 2000; Öhmanns Superpositionsmodell - Öhmann 1967; das Tilt Model - Taylor 2000) überprüft und beschlossen, dass das PaIntE Paradigma (Möhler 1998, 2001) die instrumentalen Anforderungen der aktuellen Studie am besten erfüllt. Sein Hauptvorteil gegenüber anderen *F0*-Approximationsmodellen ist die Tatsache, dass nur das feine phonetische Detail, bestimmt von einem bestimmten Segment (Phonem), von der Kontur entfernt ist. Das macht eine geglättete Kontur sogar akzeptabler perzeptiv als die Originalkontur, laut Möhlers Perzeptionssexperimente. Abgesehen von diesem Vorteil ist es auch möglich, das PaIntE Modell mit einzelnen ToBI-Labels zu verbinden.

PaIntE beschreibt eine *F0*-Kurve durch sechs Grundparameter:

- ***d*** - die Amplitude des *F0*-Gipfels in Hertz;
- ***b*** - die Ausrichtung des *F0*-Gipfels – ein relatives Maß zwischen -1 und 2, das die Position der Spitze innerhalb eines drei Silben-Fensters bestimmt;
- ***a1*** und ***a2*** - die relative Geschwindigkeit der steigenden und fallenden Sigmoiden einer *F0*-Kurve;
- ***c1*** und ***c2*** – die Amplituden der steigenden und fallenden Sigmoiden einer *F0*-Kurve in Hertz.

Während der Datenanalyse untersuchten wir Intonation im Hinblick auf die ToBI-Labels und PaIntE Parameter zugewiesen zu einer *F0*-Kurve.

Der theoretische Aspekt

Aus der theoretischen Perspektive erkundeten wir die wichtigsten SLA und allgemeinen linguistischen Theorien, um unsere ursprüngliche Vermutung über die mögliche Korrelation zwischen intonatorischer Variation und Aussprachetalent zu beweisen oder zu widerlegen. Wir wollten auch herausfinden, welche Theorien die beste Erklärung für die beiden Phänomene in Frage bieten.
Für die allgemeinlinguistischen Theorien sind wird zum Ergebnis gekommen, dass die Exemplartheorie (z.B. Pierrehumbert 1986; Nosofsky 2001; Bybee 2005; Wade et al. 2010) intonatorische Variation am genauesten vorhersagen kann, aufgrund der Tatsache, dass sie konkrete Sprachereignisse, oder Exemplare berücksichtigt, so dass die Beschreibungen der Intonationskategorien realistischer werden. Die in den Kapiteln 6-7 beschriebene experimentelle Studie scheint dieser Annahme nicht zu widersprechen. Allerdings impliziert das Ergebnis der aktuellen Untersuchung, dass die Exemplartheorie vielleicht besser für die Beschreibung von infrequenten Kategorien als von frequenten Einheiten geeignet ist, aufgrund der Instabilität der ehemaligen Art der phonologischen Einheiten.

So haben wir bei den Pitch-Akzente L* sowie den H- und L- Grenztönen, die als infrequent in der aktuellen Studie registriert wurden, keine klaren Distributionsmuster entdeckt.

Bei der Analyse der Leistung in den sechs Aussprachetalentgruppen haben wir auch festgestellt, dass die Proben größerer Gruppen (häufige Kategorien) relativ stabil im Bezug auf Variabilität und Häufigkeit sind, während Gruppen bestehend nur aus einer Person (seltene Kategorien) oft von den allgemeinen Trends auf eine inkonsistente Weise abweichen. So war die Anzahl der Tokens erzeugt von der überdurchschnittlichen weiblichen Informantin und dem unterdurchschnittlichen männlichen Probanden kleiner als in den übrigen Gruppen, was einen geringeren Grad der Variabilität voraussagt. Jedoch waren die Proben dieser kleinen Gruppen in verschiedenen experimentellen Konditionen mehr, weniger sowie gleich so variabel als die Produktionen anderer Gruppen.

Wenn der Variabilitätsgrad in einer seltenen Kategorie hoch ist, können individuelle Tokens mit einem erheblichen Abstand voneinander angeordnet sein, so dass sie sich im gleichen oder sogar in einem größeren Bereich der Werte befinden als die Tokens in einer frequenten Kategorie mit einer dichteren Bevölkerung von Tokens. Die Untersuchung von unfrequenten Kategorien im Bezug auf deren einzelne Exemplare könnte deswegen die Funktionsweise dieser Spracheinheiten aufklären, da die Verallgemeinerung in Wertbereiche ziemlich problematisch ist.

Wir haben auch weitere Fälle vom idiosynkratischen Verhalten innerhalb der Einzelpersonengruppen auf der phonologischen Ebene entdeckt, insbesondere in den Realisierungen des unterdurchschnittlichen männlichen Probanden. Er verwendete den steigenden Pitch-Akzent L*H und den hohen Grenzton H-H% am seltensten von allen weiblichen und männlichen deutschen Gruppen. Im Gegenteil, der Anteil der H* Intonationseinheiten in den Samples vom unterdurchschnittlichen männlichen Sprecher übertraf die Werte in den anderen Gruppen.

Offensichtlich weisen die oben genannten Beispiele auf die Instabilität der sehr kleinen Gruppen (seltene menschliche Kategorien), die einerseits der allgemeineren Probe entsprechen können oder andererseits von Tendenzen auf eine unberechenbare Weise abweichen können. Solche Ausreißer können mit Entwicklung einer Kategorie regelmäßige Mitglieder dieser Kategorie werden. Diesbezüglich scheint die Exemplar Theorie am besten geeignet für die Beschreibung von seltenen Kategorien und ihrer Entwicklung, weil dieses Paradigma einzelne Sprechereignisse, Exemplare, betrachtet.

Strukturelle Phonologie (z. B. Saussure 1916; Bloomfield 1933; Trubetzkoy 1958) eignet sich hingegen zur Erklärung der Mechanismen hinter den frequenten Kategorien. Sobald eine häufige Kategorie eingerichtet ist, so bevölkern ihre Mitglieder einen gemeinsamen Raum. Nach dem strukturell-phonologischen Ansatz, kann man in diesem Raum Grenzen festlegen, innerhalb derer Exemplare frei hergestellt werden können. Allerdings würde ein Exemplar, das außerhalb der Grenzen der Kategorie im Laufe deren Entwicklung entstanden ist, nicht kategorisiert. Der strukturelle Ansatz ist dabei besser prädisponiert für synchronische Beschreibungen der frequenten Kategorien, während die Exemplar Theorie auch den diachronischen Aspekt berücksichtigt.

Nachdem wir die *SLA Theorien* in Abschnitt 2.2 überprüft haben, fanden wir auch Beweise für die entsprechenden zugrunde liegenden Ideen in der aktuellen Untersuchung.

Einige der experimentellen Ergebnisse unterstützten die *Basic Variety Theorie* (Klein & Perdue 1997).Laut BV ist SLA durch eine allmähliche Steigerung des Variabilitätsgrads charakterisiert, da die Lernenden mehr Kenntnisse erwerben und so sich allmählich dem Variationsgrad der Zielsprache nähern, d.h. der muttersprachlichen Leistung.

Die aktuelle Untersuchung bestätigt teilweise diese Idee auf der phonetischen Ebene. Tatsächlich war der Grad der phonetischen Variabilität nur höher in den zwei weiblichen Gruppen, die als talentiert und sehr talentiert klassifiziert wurden (z.B. in den *c1* Parameterwerten des Pitch-Akzents L*H aufgeteilt nach Faktoren des Sprachtyps und Geschlechts - siehe Kapitel 7-8 für eine vollständige Übersicht). Diese Leistung der begabten Probanden war vielleicht der einzige Beweis zugunsten der Theorie von Basic Variety. Die anderen Ergebnisse haben zu gegenseitigen Schlussfolgerungen geführt. Außerdem wurde die Tendenz der talentierten Testpersonen zu einem höheren Variabilitätsgrad durch die männlichen Proben nicht bestätigt.

Ein weiteres Gegenargument gegen die BV ist das Fehlen der schrittweisen Erhöhung des Grades der Variation mit dem steigenden Niveau der Aussprachebegabung. Innerhalb der weiblichen Probanden wurden nur die talentierten und die sehr talentierten Informanten durch variablere Samples gekennzeichnet.

Die *Interlanguage Theorie* von Selinker (1972) und ihre wichtigsten Konzepte stellen eine relativ solide Erklärung der SLA-Prozesse vor. Einer der wichtigsten Begriffe für die gegebene Studie ist die Übertragung, oder Transfer. In der aktuellen Studie haben wir manche Instanzen vom Transfer von der L1 in die L2 entdeckt, überwiegend in den Intonationskonturen mit einer hohen finalen Tonhöhe / *F0*. Die steigenden Töne L*H und die hohen Grenztöne H-H% hatten eine größere Häufigkeit des Auftretens in den Produktionen der deutschen Muttersprachler. Dies schien durch Transfer verursacht werden, weil diese ToBI Kategorien im Deutschen weiter verbreitet sind (z. B. Mayer 1995) als in der englischer Sprache.

Diese Tendenz war jedoch von talentierten und sehr talentierten weiblichen Probanden unterdrückt. Beim Englischsprechen, reduzierten diese Informanten die Anzahl der steigenden Akzente L*H und hohen Grenztöne H-H%. Auf diese Weise näherten sich die talentierten deutschen Sprecher den englischen Muttersprachlern in Leistung an. Doch unterstützen die männlichen Proben der talentierten Sprecher nicht konsistent die gegebene Tendenz. Wir beobachteten also einen Transfer von einer typischen deutschen Kategorie in den meisten weiblichen Gruppen in die L2. Die talentierten und sehr talentierten weiblichen Probanden widerstanden dieser Tendenz, die auch zur Anpassung an das Distributionsmodell der L2 und die Muttersprachlerleistung beitrug.

Das *Diffusion Modell* (Gatbonton 1978) postuliert, dass die neu erworbenen L2-Exemplare mit L1-Tokens in freier Variation koexistieren und später in die entsprechenden Kategorien sortiert werden. Wir konnten diese allgemeine These mit den experimentellen Ergebnissen auf der phonetischen Ebene der Analyse nicht widerlegen. Wie bereits in der Interpretation der Theorie von Basic Variety erwähnt, konnten wir nur die talentierten und sehr talentierten Informanten wegen ihrem höheren Variationsgrad herausheben.
Einige SLA-Prozesse werden im *Perceptual Assimilation Modell* (Best 1995) erklärt. In dieser Theorie behauptet Best, dass in typologisch ähnlichen Sprachen eine L2-Kategorie mit dem entsprechenden L1-Sprachereignis zu fusionieren tendiert. Im Englischen und Deutschen, die oft als typologisch ähnliche Sprachen im Bezug auf Intonation bezeichnet werden (z.B. Kuhlmann 1952, Schubiger 1965, Esser 1978, Scuffil 1982, Grabe 2000), könnten die Kategorien L*H und

H-H% in den meisten Gruppen von der L1 in die L2 assimiliert werden, aufgrund ihrer konsistenten Übertragung ins Englische. Allerdings ist ein Perzeptionsexperiment notwendig, um diese Aussage über Assimilation zu beweisen.

Die *Second Language Speech Learning Theorie* (Flege 1995) expliziert den Grund für L1-Transfer in die L2. In Fleges Meinung behindert die typologische Ähnlichkeit die Diskriminierung von Kategorien. Daher werden die L1 / L2 Kategorien undiskriminiert zwischen den zwei Sprachen übertragen, ebenso wie zum Beispiel das steigende Pitch-Akzent L*H und der hohe Grenzton H-H% in den Realisierungen der meisten Talentgruppen in der aktuellen Studie.

Im Allgemeinen können die SLA Theorien einige oder mehrere Aspekte der SLA-Prozesse erklären. Allerdings scheint die Interlanguage Theorie, am besten die intonatorische Variation vorherzusagen.

Der experimentelle Aspekt

Einer der wichtigsten Teile der aktuellen Untersuchung bestand in der Validierung oder Ablehnung der Hypothesen, die im Kapitel 3 vorgestellt wurden. Alle diesen Hypothesen stammen aus der zugrunde liegenden Idee über einen Zusammenhang zwischen Begabung zur Aussprache und intonatorischer Variation. Die in dieser Studie vorgestellten Hypothesen weisen genauer auf die untersuchten Aussprachetalent- / Kompetenz-Gruppen. Wir werden kurz über diese Hypothesen und die experimentellen Ergebnisse berichten.

In *der ersten Hypothese* wird vermutet, dass die L2-Rede von weniger talentierten Sprechern weniger Variabilität in einer fremden Sprache zeigen soll. In den experimentellen Daten hatten die Produktionen der talentierten und sehr talentierten Informanten, insbesondere die der weiblichen, einen höheren Variationsgrad als die übrigen Gruppen (z.B. in den *c1* Parameterwerten der L*H Akzente gegliedert durch die Faktoren Sprachtyp und Geschlecht – siehe Kapitel 7-8 für eine vollständige Übersicht). Daher konnten wir die erste Hypothese, zum größten Teil, validieren.

In *der zweiten Hypothese* wird postuliert, dass talentierte Menschen mehr Variabilität in ihrer Rede zeigen sollen, weil sie mehr L2 Exemplare gelernt haben. So sollen sie eine größere Zahl der Exemplare gebrauchen als die weniger talentierten Sprecher. Wie wir schon für die erste Hypothese festgestellt haben, hatten die zwei talentierten weiblichen Gruppen mehr phonetische Variation in ihren Proben als die anderen Probanden.

Die *dritte Hypothese* war eine logische Folge der zweiten Hypothese. So erwarteten wir von durchschnittlichen Sprechern auch einen durchschnittlichen Grad von Variabilität, weniger variabel als die Produktionen der begabten Sprecher und mit mehr Variation als zwischen den unterdurchschnittlichen Sprechern. Diese Idee wurde teilweise durch die experimentellen Ergebnisse auf der phonetischen Ebene abgelehnt: die durchschnittlichen Informanten, vor allem die weiblichen, zeigten gleiche Variabilität wie die talentierten Gruppen in mehreren Bedingungen (z.B. Parameter *d* von L*H Akzente, abhängig vom Sprachtyp, in der Rede des durchschnittlichen Informanten in dem gesamten experimentellen Korpus in L1 und L2 und in den gelesenen Aufgaben auf Englisch – siehe Kapitel 7-8 für eine vollständige Übersicht).

Die *vierte Hypothese* wurde auch mit durchschnittlichen Sprechern verbunden. Wir behaupteten dabei, dass man die durchschnittlichen und überdurchschnittlichen Sprecher im Bezug auf Variabilitätsgrad nur schwer unterscheiden kann. Wie durch die Daten bestätigt wurde, konnte man wegen Variation nur die talentierten weiblichen Sprecherinnen separat stellen. Das Variabilitätsverhältnis zwischen den restlichen Gruppen war nicht konsistent. Daher erwies sich die gegebene Hypothese als bestätigt für die meisten Gruppen, nur mit der Ausnahme der talentierten Gruppen.

Die letzte Hypothese, *die Hypothese fünf*, postuliert, dass sich Sprecher mit höherem Grad der Aussprachebegabung die Muttersprachler in Variationsmustern nähern sollen. Der durchschnittliche sowie unterdurchschnittliche Grad des Aussprachetalents soll folglich mit einem niedrigeren Variabilitätsgrad korrelieren, und somit sollen solche Sprecher von dem L2 Modell / Muttersprachlerleistung weiter entfernt sein. Der erste Teil der Hypothese wird dadurch unterstützt, dass die talentierten und sehr talentierten weiblichen Probanden mehr phonetische Variabilität in ihren Samples gezeigt hatten als die anderen deutschen weiblichen Gruppen (z.B. den *d* Parameterwerten der H* Einheiten, gegliedert durch Geschlecht und auch in der weiblichen gelesenen Sprache in den beiden untersuchten Sprachen - siehe Kapitel 7-8 für eine vollständige Übersicht).

Als allgemeine Anmerkung im Bezug auf die Validierung der Hypothesen sollen wir feststellen, dass die experimentellen Nachweise nur die zweite Hypothese verworfen haben, während die übrigen unterstützt wurden. Allerdings sind diese Ergebnisse nur auf die phonetische Ebene bezogen. Die phonologische Analyse schien eine andere Logik zu haben. Unten behandeln wir deswegen die Ergebnisse der phonologischen Untersuchung, sowie alle zusätzlichen phonetischen Fragen.

Zusätzliche experimentelle phonetische und phonologische Aspekte

Die Analyse der ToBI Pitch-Akzente und Grenztöne hat uns ermöglicht, den Variabilitätsgrad innerhalb der Aussprachetalentgruppen auf die L1 / L2 Phänomene der Akkommodation und Transfer zu beziehen. Das Ergebnis wies darauf hin, dass die meisten deutschsprachigen Gruppen ihr L1 Verhalten in die englische Sprache übertrugen, während talentierte Sprecher versuchten, sich dem L2-Modell anzunähern. Diese Tendenz war besonders ausgeprägt in den L*H Pitch-Akzenten und H-H% Grenztönen. Die talentierten Informanten reduzierten die Prozentsätze dieser typisch deutschen Kategorien in ihrer L2 Rede. Im Deutschen hingegen sind diese diese Spracheinheiten bedeutend weiter verbreitet. So waren die Produktionen der talentierten Sprecher näher an den Realisierungen von englischen Muttersprachlern.

Wie die experimentellen Ergebnisse beweisen, folgen die phonetische und die phonologische Ebene einer unterschiedlichen Logik im Bezug auf den Variabilitätsgrad in den einzelnen Pitch-Akzenten und Grenztönen. Die phonetischen Ergebnisse beziehen sich auf den quantitativen Aspekt der Variabilität – mehr / weniger Variation, und die phonologische Analyse spiegelt vielmehr die Interlanguage-Phänomene vom Transfer und Akkommodation in der Intonation wider. Die Analyse des Korpus auf der Textebene schien noch ein anderes Produktionsmuster über die Aussprachebegabungsgruppen vorzuschlagen.

Während der globalen phonologischen Untersuchung auf der Textebene verglichen wir ToBI Transkriptionen der beiden Texte unter Analyse der Fabel "Der Nordwind und die Sonne" und des Auszugs von "Mrs.McWilliams" von Mark Twain, um die Einheitlichkeit / Variabilität der prosodischen Interpretation der Texte zu untersuchen. Die erste Tendenz der deutschen Muttersprachler in L1 bestand darin, dass die Transkriptionen einheitlicher waren als in allen ihren englischen Produktionen, die mehr Variabilität und folglich weniger Einigkeit zwischen den Sprechern zeigten. Diese Tatsache könnte dadurch erklärt werden, dass sich die Sprecher mit ihrer Muttersprache wohler und ruhiger fühlten. Deswegen wurde ihre Interpretation der experimentellen Texte neutraler und einheitlicher, so dass auch bestimmte Produktionsmuster in einer Gemeinschaft von Referenten entstehen können.

In den englischen Versionen der beiden untersuchten Texte entdeckten wir auch einige konsistente Muster der prosodischen Textinterpretation in den untersuchten Gruppen. So zeigten die englischen Muttersprachler sowie die durchschnittlichen Sprecher den höchsten Grad der Vereinbarung zwischen den Sprechern, unabhängig vom Geschlecht in den Proben von "Der Nordwind und die Sonne" als auch in weiblichen Realisierungen von "Mrs.McWilliams". Eine Erklärung für solche Ergebnisse zwischen diesen Probanden wäre eine relativ neutrale Haltung zu der Aufgabe, die auch neutrale Realisierungen verursachte.

Die sehr talentierten und untalentierten Sprecher demonstrierten hingegen den höchsten Grad der Variabilität in prosodischer Textinterpretation, was besonders deutlich in ihren Produktionen von "Mrs.McWilliams" ausgeprägt war. Wir würden verschiedene Erklärungen für diese Tatsache für die beiden Gruppen vorschlagen. Die talentierten Informanten strebten vielleicht auf eine abwechslungsreiche und "interessante" Interpretation der experimentellen Texte und hatten dabei unterschiedliche Intonationsmittel verwendet. Dies führte zu einem höheren Grad der Uneinigkeit zwischen den Sprechern. Die untalentierten Informanten, auf der anderen Seite, mit ihrem niedrigeren Leistungsniveau, waren vielleicht nicht sicher, welche Intonationseinheiten im gegebenen Kontext einzusetzen waren, was auch größere Uneinigkeit innerhalb der Gruppe in dieser Hinsicht verursachte.

Die Gründe für größere / kleinere Grade der Einigkeit oder Variabilität in prosodischer Textinterpretation in den Gruppen waren offenbar nicht nur von den phonetischen oder phonologischen Faktoren geführt. Eine Rolle spielten vermutlich auch die außersprachliche Prädisposition eines Sprechers zu der experimentellen Aufgabe (neutral – für durchschnittliche deutsche Sprecher und englische Muttersprachler; mehr beteiligt - für die talentierte Gruppe) sowie deren Leistungsstand im Englischen (die untalentierten Probanden besaßen vielleicht nicht ausreichend Exemplare in ihrem Intonationsinventar). Somit wird der Variabilitätsgrad dadurch bestimmt, dass ein Sprecher ein ausreichendes Inventar an Sprachmitteln besitzen soll. Man soll auch wissen, in welchen Kontexten diese Mittel zu gebrauchen ist. Eine persönliche Einstellung ist natürlich keine eindeutige Bedingung, obwohl eine positive Einstellung zu mehr sprachlicher Variation zu motivieren scheint. Die untalentierten Sprecher wissen nicht, wie sie Spracheinheiten richtig gebrauchen sollen, sogar wenn sie gut motiviert sind, und so wenden sie Sprachmittel inkorrekt und unterschiedlich an, was auch größere Uneinigkeit zwischen den Sprechern verursacht.

Es ist auch eine zusätzliche phonetische Tendenz im Hinblick auf die experimentellen Hypothesen sowie die allgemeinen sprachlichen und SLA Theorien entstanden: talentierte und sehr talentierte weibliche Sprecher haben ihre PaIntE Parameter auf einer höheren mittleren *F0*-Niveau realisiert als die anderen Gruppen in einer Reihe von experimentellen Bedingungen (siehe Abschnitt 7.2.1 für eine Zusammenfassung). Dieser Befund scheint mit der oberen Tendenz der talentierten Sprecher für einen höheren Variabilitätsgrad in prosodischer Textinterpretation verbunden zu sein. Die Probanden haben wahrscheinlich versucht, nicht nur phonologisch ihre Rede variabler zu machen, sondern erhöhten auch ihre *F0*, sogar höher als die Muttersprachler. Frühere Studien berichteten auch über ein solches Phänomen der „Overhitting“ des L2-Targets (z.B. Bohn & Flege, 1992).

Zum Abschluss der Zusammenfassung der experimentellen Ergebnisse ist es wert zu erwähnen, dass einige der oben erwähnten Tendenzen mit unseren früheren Untersuchungen über die gleichen Gruppen von Testpersonen übereinstimmen. So haben wir bisher einen höheren Variabilitätsgrad sowie eine höhere Tonlage innerhalb der talentierten Sprecher, insbesondere den weiblichen betrachtet (Anufryk 2009; Anufryk & Dogil 2009). Wir fanden auch Beweise in unseren früheren Untersuchungen für die Tendenz dieser Gruppen, die Anzahl der steigenden Pitch-Akzente und hohen Grenztöne in der L2 unterzubringen (Anufryk 2008; Anufryk et al 2008;. Anufryk 2009; Anufryk & Dogil 2009). Für die intonatorische Textinterpretation waren unsere bisherigen Ergebnisse auch ähnlich: weniger Variabilität in allen deutschsprachigen Proben in ihrer deutschen Muttersprache; auf Englisch mehr Einigkeit zwischen

Muttersprachlern und der durchschnittlichen Gruppe und mehr Uneinigkeit zwischen den talentierten sowie untalentierten Sprechern (Anufryk 2009).

Schlussfolgerungen

Vom linguistischen Standpunkt lag der Fokus unserer Aufmerksamkeit auf den häufigen phonologischen Kategorien, die offenbar mehr mit den Konzepten der Exemplartheorie and struktureller Phonologie übereinstimmen. Wir haben festgestellt, dass Exemplartheorie für eine dynamische Beschreibung der Kategorien in ihrer Bildung am besten geeignet ist, da sie sich auf die individuellen phonetischen Repräsentationen konzentriert. Der strukturelle phonologische Ansatz beschreibt Kategorien statisch: alle ausreißenden Instanzen können nicht kategorisiert werden.

Die unfrequenten Spracheinheiten wurden in der aktuellen Studie nicht behandelt. Unserer Meinung nach bieten seltene Klassen jedoch einen Einblick in die diachronische Perspektive, d.h. die Entwicklung der Kategorien. Sie liefern auch gleichzeitig synchronische Informationen darüber, wie sie zu einem gegebenen Zeitpunkt funktionieren. Es wäre daher interessant, die Aufmerksamkeit auf seltene Kategorien in unserer zukünftigen Arbeit zu widmen.

Von der SLA Perspektive, waren manche Ergebnisse der gegebenen experimentellen Untersuchung im Vorhersagen von den großen SLA Theorien erklärt - Basic Variety, Interlanguage, Diffusion Modell, Perceptual Assimilation Modell und Second Language Speech Learning Theorie. Die gegebene Studie konnte jedoch zum Verständnis der ihnen zugrunde liegenden Prinzipien in einigen Aspekten beitragen – z.B. das Fehlen einer schrittweisen Erhöhung der Intonation Variabilität mit steigenden Aussprachebegabung / Kompetenz, im Gegensatz zu den Vorhersagen der Basic Variety Theorie. Die Methodik der aktuellen Studie hat uns ermöglicht, detaillierte phonologische Informationen über einzelne Intonationsereignisse durch die ToBI-Labels zu erhalten. Die sechs PaIntE Parameter lieferten wiederum phonetische Details. Abgesehen davon konnten wir den außersprachlichen Faktor der Aussprachebegabung durch ihre Manifestation in Kompetenz untersuchen.

Wenn wir weiterhin ein Experiment mit unserer Methodik durchgeführt hatten, das auf den theoretischen Prämissen der SLA und allgemeinlinguistischen Theorien basiert wurde, konnten wir Unterstützung für die meisten der experimentellen Hypothesen in unseren Daten finden sowie auch zusätzliche experimentelle und theoretische Fragen beantworten.

Chapter 1

Aptitude and other individual differences in foreign language learning and acquisition

The current investigation is not concerned directly with testing and measuring the foreign language learning aptitude. Nevertheless the human ability for pronunciation, in particular, prosody, has been the main inspiration for this study. It is therefore important to define the notion of pronunciation aptitude, delimit it from the other individual differences, to which it is sometimes equalled, as well as draw implications for further research.

The majority of studies on language aptitude have focused on one or several of the following, rather contradictory issues:

- testing aptitude
- correlating aptitude with other individual and external variables
- studying the biological predisposition for language aptitude
- investigating the "outcome" of aptitude, i.e. an individual's achievement \ proficiency in a given language

While it is not possible to draw a direct line between these domains, they do deal with important aspects of language aptitude. Thus, we will attempt to outline the main ideas relevant for the current investigation in the sections that follow.

1.1 Aptitude tests

Arguably, the most influential contribution to the study and measurement of foreign language aptitude to date has been rendered by J.B. Carroll, who created the first comprehensive test battery – the MLAT (Modern Language Aptitude Test), inspiring the multitude of studies and tests that followed. He investigated the phenomenon in question only in terms of supervised or unsupervised language instruction, aimed at performing tasks and programs. Following that approach, he defined foreign language aptitude as "*current state or capability of learning that task – if the individual is motivated and has the opportunity of doing so*." (Carroll 1981: 84) This focus on solely instructional influence on foreign language aptitude was further criticized due to the fact that ability plays an important part in naturalistic L2 acquisition to an equal degree, as it appears to be an equally great challenge to form target-language categories from mere exposure to them without explicit instruction, which, in turn, requires ability on the part of the learner (Skehan 2002).

Carroll outlined the four components of foreign language aptitude:

1) phonetic coding ability;
2) grammatical sensitivity;
3) rote learning ability – the ability to associate sound and meaning;
4) inductive language learning ability – the ability to infer rules, given some language materials (Carroll 1962).

The importance of Carroll's work for the study of pronunciation aptitude consists in his recognition of this phenomenon as a separate faculty, in which mimicry, or imitation, of sounds and stretches of sounds plays an important part. Apart from that, Carroll intended to include tasks for the "*control in imitating stress and intonation*" (Carroll 1953: 194). However, tests of prosodic ability were not part of the classical test.

The later aptitude tests relied on Carroll's basic four-component notion of foreign language aptitude, stressing this or that component to a greater degree than the MLAT. Pimsleur's Language Aptitude Battery (Pimsleur 1966), the PLAB, emphasized the auditory component and the inductive ability, which was not tested to a sufficient degree in the MLAT; the Defense Language Aptitude Battery (Petersen & Al-Haik 1976), or the DLAB, tested for the highest level of inductive ability through audiovisual materials. Unlike the previous tests, which investigated the phonetic coding ability, the VORD (Parry & Child 1990), named so after an artificially created Turkish language, only focused on syntax and morphology. The latest of the best-known tests batteries, the LLAMA, was also largely based on the MLAT, and, as the creators state, it "*should not be considered a replacement for MLAT in high-stakes situations*" (Meara 2005). The former has the same basic structure as the Modern Language Aptitude Test, and it investigates the phonetic coding ability by means of memorizing words of an unknown artificial language and through sound-symbol correspondences.

The only test departing from the MLAT was the Cognitive Ability for Novelty in Acquisition of Language – the CANAL-F (Grigorenko, Sternberg & Ehrman 2000). Its differences from the previous tests can be summarized as follows: a) it stressed the novelty of material; b) it used naturalistic situations; c) it assessed an individual's ability during the course of the test; d) it both assessed a learner's aptitude and produced the appropriate teaching and learning strategies; e) it was an adaptive test, i.e. it adjusted to a learner's responses. However, the test's main focus was limited to syntax and morphology.

As seen from the overview provided, most of the language aptitude tests shared the same basic structure introduced in the MLAT, or reduced it to the syntactic and morphological components. The phonetic ability was only tested by imitating, memorizing or learning the sound-to-letter correspondences of lexical items. However, pronunciation ability cannot be reduced to imitation and memorization alone – dynamic production is of equal importance. Apart from that, prosodic ability, which was not tested at all in those batteries, constitutes an important faculty within the general phonetic ability.

The fact that language aptitude tests were so limited in the inventory of skills / abilities they investigated led to a considerable amount of criticism to their validity. This, in turn, coupled with other factors, led some researchers to question the very existence of language aptitude.

1.2 Aptitude or something else?

Language aptitude, understood here as an individual's potential for learning as well as the natural acquisition of a foreign language is a phenomenon, which is difficult to measure and adequately explain. Even with the current advances in its investigation, it still remains for the most part an enigma. Therefore, some researchers regard it as inapplicable to the learning process. Marinova-Todd, for example, stated that if aptitude were to account for the success or failure in foreign language learning, it would not be satisfactory, as neither teachers nor students know how to influence it (Marinova-Todd 2005).

One might provide several counterarguments to this idea. First of all, the knowledge of an individual's abilities can be used for creating teaching and learning materials best suited for this or that person, as has been done, for example, in an investigation by Wesche (1981). Apart from that, even without changing the general teaching paradigm, knowing a learner's potential is beneficial in a way that a teacher might use more individualized approaches to students, if he or she is aware of a learner's "limits" and might therefore motivate further efforts on the part of the latter. Thus, boosting inherent individual abilities through external influences can be applied to facilitate learning.

Some researchers tend to equal those favourable external factors, i.e. exposure to the target language, qualified instruction etc., alongside with a number of internal characteristics, to the aptitude per se (e.g. Bialystok 2002). Other scholars only regard aptitude as relevant for instructed learning as opposed to natural acquisition (Carroll 1981; Gardner 1985).

In most learning cases, successful L2 attainment is only possible as an outcome of several favourable conditions / factors. Further on, some of those conditions / factors correlate with language aptitude to a considerable degree, which has been shown in a multitude of studies. However, the importance of foreign language learning aptitude cannot be ruled out, as none of these factors is fully able explain the differences in SLA between various speakers, especially if all the other internal and external characteristics are equal.

1.3 Aptitude versus other individual and external variables

The following discussion of an interrelation between foreign language learning aptitude and other individual differences, as well as external factors, is not intended as a comprehensive insight into an L2 learner's psychology and neurocognitive processes. For insightful overviews on cognitive variables (empathy, mental flexibility, working memory and intelligence), personality, musical ability and the neural predisposition for language aptitude, it is advisable to consult the recent volume edited by Dogil & Reiterer (2009). In the current study we will only explore the individual variables most easily and frequently correlated with successful attainment in L2 and language aptitude, i.e. those of age and motivation. Apart from that, exposure will be explored as an important external factor for L2 acquisition. Finally, we will attempt to formulate the relationship between aptitude and proficiency, as it constitutes a crucial starting point for the current investigation.

1.3.1 Age

The problem of age and its influence on a learner's language abilities and subsequent performance became one of the central ones in the 60s, when Eric Lenneberg (1964, 1967) introduced the idea of a critical period for language acquisition. According to his hypothesis, due to the loss of cerebral plasticity, post-pubertal acquisition of language can only result from a conscious effort on the part of the learner. Even this laborious process cannot help overcome a foreign accent. Younger learners, on the other hand, acquire languages through mere exposure to them, and they therefore have an advantage over adults and teenagers in this respect.

Although Lenneberg's original research concerned children recovering L1 abilities after aphasia, his seminal work was then applied to second language acquisition.

This cornerstone hypothesis inspired a multitude of studies in the following decades on a variety of aspects. Many of those investigations are still subject to debate. First of all, there exists no consensus as to the actual starting point of maturational constraints. The majority of studies proposed the age of 10-12 years (e.g. Lenneberg 1967; Seliger 1978; Scovel 1988; Johnson & Newport 1989; Long 1990; Pinker 1994), when the lateralization process is completed, i.e. when specific language functions are assigned to the left and the right hemispheres. However, as pointed out by Walsh and Diller (1981), the critical period may start earlier, mainly by age 6-8, with regard to the formation of new phonetic categories due to the consolidation of pyramidal axon connections. Secondly, various studies viewed those ages as either the starting or the end point of the neurological changes leading to the decline in language performance (Birdsong 1992).

Apart from that, a distinction was often made on the critical period for the acquisition of the phonological system versus the acquisition of syntax and morphology. Thus, the above-mentioned critical periods should mainly influence the acquisition of phonology, i.e. ages 6-12, whereas a naturalistic acquisition of syntax and morphology should decline after the age of 15 (e.g. Long 1990). This distinction was linked to the fact that the phonological component constitutes a "lower-order" function which pertains to the sensory-motor domain and therefore ceases to develop after puberty. Syntax and morphology, on the other hand, are "higher-order processes" in the domain of cognitive development located in the cortical regions, in which neural functions develop throughout an individual's life in response to linguistic input. It made some researchers logically conclude that there are in fact "multiple critical periods" of language acquisition (Seliger 1978, Long 1990).

To complicate the problem further, some studies provided evidence that an early start in learning a language does not guarantee the actual success and higher proficiency levels. Indeed, some learners retain a foreign accent despite the early exposure to a given L2 (Flynn & Manuel 1991; Flege et al. 1997; Pallier et al. 1997; Abu-Rabia & Kehat 2004; Moyer 2004; Trofimovich 2005).

At the opposite end of the spectrum, there exist late learners who are capable of achieving native-like proficiency despite the post-pubertal start in L2 acquisition. An array of studies on ultimate attainment in SLA reported about those exceptional learners of various languages: German (Moyer 1999, 2004); Hebrew (Abu-Rabia & Kehat 2004); English, German and French (Bongaerts 1999; Bongaerts et al. 2000); Arabic (Ioup et al.1994) – to name a few examples.

The controversies of the critical period hypothesis, mainly, the absence of consensus on the actual time frame when maturational constraints come into effect for the acquisition of various aspects of language, as well as differences in performance between early and late starters, implies that there is a great deal of individual variation in the age factor.

Flege, speculating on these controversies, suggested that other factors might be at play when tackling the issue of adult versus child foreign language acquisition. Among those other factors he mentioned, first of all, the different modes of L2 processing at different ages. More specifically, due to a greater degree of L1 exposure, an adult's phonetic categories are much more firmly established than those of a child. For that reason, when being confronted with a new learning experience, e.g. a new sound, an adult is more likely to classify this new instance as a member of an existing L1 category. The above is typical of the so-called "phonetic" learning mode. Children, on the other hand, form and perceive new L2 categories more quickly and easily because their L1 units are still flexible and not that firmly established, which is characteristic of the "auditory" learning mode, according to the author (Flege 1987). Other studies named the following factors contributing to the child-adult differences in L2 acquisition:

- quality and quantity of L2 input – children are usually exposed to simpler language than adults, who use it more for describing abstract notions and states and are therefore less fluent in the "down-to-earth" speech (Cochrane 1977; Burling 1981);
- fear and anxiety – adults are more afraid to make a mistake than children (Guiora et al. 1972; Schumann 1976, 1978);
- short-term vs. long-term learning effects – adults outperform children in the former kind of learning, whereas children are better in the long-term (Snow & Hoefnagel-Hoehle 1977, 1978; Krashen et al. 1979);
- language structures under examination – children learn and produce phonological units more quickly and accurately (Fathman 1975), while adults and teenagers outperform younger children in syntax and morphology (Fathman 1975; Krashen 1982).

As evident from the arguments provided, age is not the only explanation for the differences between early and late foreign language acquisition and learning. Nevertheless according to the vast bulk of studies, language learning ability does decrease with age for the majority of speakers. Some individuals, however, do retain this ability throughout their lives. Such people cannot be disregarded as simple outliers, as "*for all natural populations, exceptions abound*" (Scovel 2000: 217). Moreover, these individuals represent the top of the proficiency hierarchy and constitute a considerable amount of language learners – reportedly, from 5 (Selinker 1972) to 35 per cent of all people (Seliger 1978). The latter number concerns individuals who are unencumbered by the loss of cerebral plasticity at puberty.

The above implies that up to a third of the population of learners, arguably, is potentially capable of attaining accent-free speech in a foreign language. Therefore, the talented individuals, who do possess near-native skills as such, ought to be looked at more closely to understand, which techniques and principles they are guided by in speech production. This knowledge might have a lot of further implications for the applied domain (how to instruct language learners to use the best of their inherent ability), general linguistic descriptions, as well as methodological developments in experimental design and speech technology.

Age in general, and the critical period phenomenon in particular, is one of the major factors influencing the SLA performance / proficiency. Motivation is named in a considerable number of studies as another crucial factor, which is claimed to override the maturational constraints in importance in some cases.

1.3.2 Motivation

All of the studies on successful ultimate attainment mentioned in Subsection 1.3.1 investigated highly motivated L2 subjects.

Bongaerts (1999) reported two groups of very advanced L2 learners of English and French, who attained native-like pronunciation due to, as the authors suggested, a combination of motivational and instructional factors, alongside with a significant amount of L2 exposure, which compensated for a late start in L2 learning. Moreover, the study indicated that some of these learners were consistently assessed as being native-like and, in some cases, even better than the native speaker controls, as judged by a group of native experts. In a subsequent study, Bongaerts et al. (2000) investigated a group of late learners of Dutch and obtained similar results; the only difference being the fact that L2 speakers were judged in the lower range of the native controls, as opposed to the previous study.

The two investigations carried out by Moyer (1999, 2004), who looked into the pronunciation of L2 speakers of German, also put forward the idea that a high motivational level, both professional and personal, correlates significantly with a successful L2 outcome. Nevertheless the age of onset remained the strongest predictor of L2 success in these studies.

Abu-Rabia and Kehat (2004) explained their subjects' successful performance in Hebrew through a combination of affective variables, i.e. "*mainly motivation and talent*", including *"possibly other cognitive variables as well"*.

It seems almost unquestionable that the subjects' positive motivation in the above studies did contribute to some degree to a native-like L2 performance in a given language. However, those descriptions only touched upon highly proficient L2 speakers with a high motivational level, coupled with other factors like consistent L2 input, pronunciation instruction and talent. It is therefore not clear if motivation alone could account for successful L2 outcome.

Moreover, increased motivation in those subjects could be the result of their success in a foreign language, as, indeed, some studies claimed that successful performance causes motivation (e.g. Burstall 1975; Hermann 1980; Crookes & Schmidt 1989; Skehan 1986; Weiner 1992). Sparks and colleagues also pointed out in their study that *"poor foreign-language learners are equally as motivated to learn a foreign language as good foreign-language learners but perceive themselves as having weaker skills in both the oral and the written aspects of language learning"* (Sparks et al. 1998: 183). The authors further hypothesised that the differences in L2 learners' affective states most likely stem from the differences in language skills. Support for this hypothesis can be found in several other studies on motivational influences upon L2 learning (Oller 1981; Au 1988; Long 1990).

Flege (1987), seconded by Piske et al. (2001), summed up the above findings by pointing out that no existing research up to date has come up with sufficient evidence of the fact that varying levels of motivation lead to differences in L2 success, especially with regard to pronunciation. Indeed, a high motivational level does not guarantee accent-free speech. As mentioned by Moyer, some people *"pursue such a goal [a learning goal] in spite of their own limits (e.g. limits in perceptual abilities and limited access to authentic input etc.)"* (Moyer 2004: 40).

It logically follows from the above statement that, while motivation may contribute to successful L2 attainment to some degree, it cannot override an individual's natural ability in importance, as it is the ability itself that delineates where a learner's "limits" lie. This observation is consistent with the results obtained in several studies on the interaction between aptitude and motivation (Colletta, Clement & Edwards 1983; Clement & Crudenier 1985).

As a final remark, it is worth pointing out that the influence of motivation upon L2 attainment should be accurately appreciated, and only with regard to other social and personal variables.

The same appears to hold true for another external factor, which is sometimes claimed to be the sole explanation of successful L2 performance – that of exposure.

1.3.3 Exposure

Exposure is understood in the current study as the amount of L2 input received by an individual up to a certain moment in time. A distinction can be made between naturalistic exposure, commonly equalled to the length of residence (LOR) and supervised exposure in a classroom environment.

The overview provided in the previous subsections already contains some implications as to the relationship between language learning aptitude and exposure. Thus, an earlier start in learning a given language should generally lead to a greater degree of L2 exposure, which, in turn, is claimed to have a positive influence on L2 attainment. Following this logic, younger learners should have an advantage over older individuals, which is true for the majority of a given population, due to the influence of maturational constraints. However, this is not always the case, as there are both early starters who retain a foreign accent (Flynn & Manuel 1991; Flege et al. 1997; Pallier et al. 1997; Abu-Rabia & Kehat 2004; Moyer 2004; Trofimovich 2005) and exceptional late learners with less L2 experience (Ioup et al.1994; Moyer 1999; Abu-Rabia & Kehat 2004) or even only incidental L2 exposure (Bongaerts 1999; Bongaerts et al. 2000).

Departing from the maturational factor, a wealth of studies concentrated on the two types of exposure as independent variables, or in combination with other individual factors.

The examination of the length of residence as a predictor of L2 performance has lead to conflicting findings in those investigations. Some researchers found a positive effect of LOR (e.g. Purcell & Suter 1980; Flege 1988; Flege & Liu 2000; Moyer 2004); others concluded the length of residence alone to be an insufficient predictor (Moyer 1999; McAllister 2001; Piske et al. 2001; van Dommelen & Husby 2009).

In a similar fashion, the amount of classroom-based exposure may or may not have a positive influence upon L2 performance. In particular, phonetic/phonological training was reported to have a varying effect on the outcome. Moyer (1999) reported that both overt segmental and supra-segmental training predicted a closer-to-native performance. However, no significant correlation was found with the regard to the native-level performance. A positive effect was also found for long-term phonetic training (Elliot 1995; Moyer 1999) and audio-visual instruction (DeBot 1980, 1983; DeBot & Mailfert 1982; Michas & Berry 1994). Other studies, on the other hand, report no effect (Derwing et al. 1998) or contradictory results (Dickerson & Dickerson 1977; Tarone 1988).

Indeed, phonetic instruction may have a positive effect on a learner's pronunciation. Nevertheless anecdotal and research evidence demonstrates that not all same-age learners cope with those training tasks equally, even if the amount of support and input is the same. Teachers of phonetics, including the author of the study in question, have pointed out from their experience that some learners are not able to discriminate the phonological contrasts perceptually and do not manage to produce them in an adequate way. Other learners, in contrast, seem to grasp the material almost effortlessly. Such a difference in performance, in turn, results in different grading of learners – a surface representation of academic success.

The above overview implies that the degree of L2 exposure cannot, as a rule, be the sole predictor of L2 performance. Aptitude plays an important part both in the naturalistic acquisition through the length of residence and in the classroom learning.

Support for this idea can be found in Purcell and Suter (1980), who found aptitude (oral mimicry) to be the most important individual variable following the first-language factor. Following this idea, Ross and coworkers stated that *"individual differences in language learning aptitude ... possibly indicate a factor that compensates for some learners' relatively late and infrequent access to English"* (Ross et al. 2002:295). Thus, exposure does not contribute to language aptitude: talented speakers are capable of successful attainment even despite the lack of exposure.

What exposure does contribute to is a phenomenon somewhat adjacent to aptitude – that of language proficiency.

1.4 Aptitude versus proficiency

Proficiency occupies a special position in the array of individual differences reflecting the actual outcome of all the other variables. Thus, motivation, age of learning, exposure and other neurological, cognitive, psychological as well as external factors, alongside with language aptitude, all add up to the general level of L2 proficiency.

This cause-and-effect relationship sometimes makes it difficult to distinguish between L2 learning aptitude and proficiency. We follow the distinction formulated by Carroll (1981) of language aptitude being the potential for language learning and acquisition, whereas proficiency delineates an individual's inventory of specific language skills (lexical, morphosyntactic and phonological) and the ability to use them in an appropriate context at a given point. In other words, proficiency is the result of what can be achieved given a certain degree of learning aptitude, coupled with the other variables mentioned before.

If truly aptitude had to be delimited from all the other variables, the latter should be levelled to a maximum degree. In other words, if those variables are at least comparable in a given group of individuals (the ideal situation of them being absolutely equal is almost impossible), the attained level of proficiency should reflect a speaker's degree of language aptitude.

Speakers can further be located on the aptitude – proficiency scale to study what specific language structures and speech production principles, i.e. proficiency elements, correlate with greater or lesser degrees of aptitude.

One such element proposed for investigation in the current study is intonational variation. Its relation to language aptitude, as well as the general linguistic and SLA perspectives of this phenomenon will be treated in due detail in the following chapters.

Chapter 2

Intonational variation in the light of general linguistic theories and theories of second language acquisition

The study of variation is one of the central domains in linguistics, as we find it on all levels and within all units. From one perspective, it stands for the richness of language and the ability of individuals to take advantage of it. Indeed, an individual's proficiency in a given language, be it native or foreign, is determined by his or her ability to select and employ the language means prompted by the corresponding communicative situation. While the majority of people are capable of getting their message through to the interlocutor, it is the degree of suitability, expressiveness, as well as variation, which differentiates a gifted speaker from a less successful one. Thus, the ability to vary lexical units and grammatical patterns has often been mentioned as a talent in its own right.

Is this true for intonation? It is possibly the most intricate language phenomenon in itself. Precisely because of its high variability, there exists no consensus as to the inventory of intonation categories and their functions in any given language, which only complicates cross-linguistic analysis, even within a homogeneous sample. If, on the other hand, one needs to analyze speech realizations of a non-uniform group, e.g. speakers of varying proficiency and aptitude, this contributes still an additional layer of variability. It is not clear therefore, whether one should expect more variability from more proficient or less proficient individuals in their native or foreign language.

One might, however, draw some implications from previous research on the problem. In particular, for the analysis of cross-linguistic intonational variation across speakers with a varying degree of pronunciation aptitude, three aspects are to be specified: the general idea of what language variability is, its connection to second language acquisition phenomena and the typological domain. Finally, it should be identified whether variation patterns are in any relation to an individual's proficiency level and aptitude.

In view of these considerations, we will first try to locate the phenomenon of intonational variation within the multitude of the existing linguistic paradigms and theories of second language acquisition. Further on, we will present an approach to cross-language (more specifically, English-German) intonational variation integrating the implications drawn from these theories, alongside with our general considerations on language aptitude set forth in Chapter 1, which, in turn, will serve as a conceptual premise for the subsequent experiment.

2.1 Variation in linguistic theories

The correlation between constant and variant features in the language system is a cardinal problem, which is explored in the majority of modern linguistic investigations. Despite the fact that the general idea, that a language system is a heterogeneous entity, is broadly accepted nowadays, views on language variation have undergone a considerable evolution. It has been a long way from the pure structuralist abstraction to the atomic exemplar matrix, i.e. from the language code to its dynamic speech representation.

The key to understanding language variation lies in resolving two major issues. Firstly, it is important to identify the inventory of language units and their features, as well as define variability limits for each unit. Secondly, determining the dimensions of variation is crucial: should it be restricted to the systemic level, i.e. categorical variation, or should it also account for the plurality of speech realizations within a category.

2.1.1 Structural Phonology

In Structural Phonology the systemic level clearly took the focus of attention. Language was considered to be a relatively self-contained sign system, which is stored as a whole only in the collective consciousness (Saussure 1916). For this reason, structuralists were mainly interested in the formal description of this system.

Although intonation was not the central notion in early structuralist studies, it did receive a basic description as a dichotomy of falling and rising movements, the former having the meaning of finality and the latter bearing the continuation function (Trubetzkoy 1958).

Jakobson & Halle pointed out that *“tone level, or tone modulation … are always purely relative and highly variable in their absolute magnitudes from speaker to speaker, and even from one utterance to another in the usage of the same speaker.”* (Jakobson & Halle 1971: 115) Despite this obvious acceptance of phonetic variation as a speech phenomenon, it was only seen as a source of material for the higher, phonological, analysis levels. Structural linguists were not interested in idiolects and individual realizations as such; they strived to generalize the language patterns of a certain community. Equally irrelevant was phonetic variability in cross-linguistic research and studies of second language acquisition: language comparison (Hjelmslev 1974), as well as learning (Shcherba 1974), only consisted in getting a full command of its categories and invariants, regardless of the speech peculiarities.

In the language system intonation was seen as sharing the same properties as the segmental units, cf. the terms 'secondary phonemes' (Bloomfield 1933), 'suprasegmentals' (Lehiste 1970). Prosodic variation is thus generated from an abstract set of features which may vary to a certain degree, but are still perceived as the same entity, i.e. variation is free within certain abstract boundaries. Such an understanding of phonological categories was predominant for a substantial period of time. However, what this approach failed to explain were phenomena adhering to a certain category, but lying outside the established boundaries.

With all due respect for the significance structural linguistics had for the understanding of language variation in general, it has to be stated that structuralist descriptions were still very abstract and limited to the phonological norm. However, without regard to the phonetic variation within a given category, it is impossible to get a full perspective of how language units function in speech. Even comparative analysis is hampered when identical categories in two languages differ exclusively in their distributional phonetic characteristics.

2.1.2 Generative Phonology

Generative Phonology shared the primacy of the phonological level introduced in structural studies.

In 'The Sound Pattern of English' (SPE) Chomsky & Halle (1968) made several important assumptions about intonation: namely, they introduced the nuclear stress rule – part of the transformational cycle, a process assigning special weight to the last stress in a prosodic pattern as compared to all the other stresses. This rule postulated the complexity of the prosodic contour as a multi-level structure of temporal, stress and pitch parameters, which, in its turn, predetermines its variability. SPE only pointed out its correlation to syntax, or the surface structure of the utterance.

> *Once the speaker has selected a sentence with a particular syntactic structure and certain lexical items, the choice of stress contour is not a matter subject to further independent decision. That is, we need not make a choice among various 'stress phonemes' or select one or another superfix (SPE: 298).*

Whereas SPE's general understanding of intonation contour structure complied with most modern definitions, the above idea of absolute correspondence between the syntactic and phonetic levels rejects the notion of prosodic variation whatsoever. If it were true, there would be one and only one realization for each utterance. The strongest counter-argument would most probably be presented by the syntactically ambiguous utterances, where intonational variants are absolutely indispensable for this or that semantic interpretation. Ample evidence has also been found for non-ambiguous sentences to state that the relationship between syntax and prosody is non-linear (Selkirk 1995). Furthermore, extralinguistic context, as well as the speaker's individual characteristics, adds up to the plurality of the possible prosodic realizations.

As evident from the arguments provided, Generative Phonology did not differ significantly from the structuralist positions on prosodic variation.

2.1.4 Optimality Theory

Within the optimality theoretic paradigm, prosodic variation consists in retrieving the optimal variant resulting from the interaction of the three major elements – generator, constraints and evaluator (Prince & Smolensky 1993), whereby in the general ranking schema prosodic constraints dominate the morphological ones (McCarthy & Prince 1993: 145). It occasionally happens that in the course of evaluation more than one candidate is selected as optimal, i.e. best (e.g. Selkirk 2000: 250). Thus, OT does allow of a considerable variation in prosody, though only phonologically and syntactically incurred (e.g. Féry & Truckenbrodt 2005), whereas phonetic variants remain beyond the theory.

2.1.5 Prototype Theory

The next big step in variation research was marked by Prototype Theory with Eleanor Rosch as one of its founders. The new principle of categorization consisted in singling out a centroid, the best representation of a class which is most easily recognized and most often produced by speakers.

Crucial for phonetics was the notion of the Perceptual Magnet Effect introduced by Kuhl (1991). Exploring vowel space, she found evidence that *"the prototype of the category assimilated neighbouring stimuli, effectively pulling them toward the prototype"*. That discloses the underlying psychological structure of speech categories, as the author suggested.

The objectivity of this effect was supported by some studies of intonation. Schneider & Möbius (2005) reported finding the prosodic prototype for the statement category in German, which is characterized by very low *F0* values. However, no Perceptual Magnet Effect was registered for the question category in that investigation.

Prototypicality effects may occur in the perception of some intonational categories, but no studies have dwelled upon the production of prosody in that respect. Indeed, if speakers only produced prototypes, intonation would be fairly uniform. Yet, it varies greatly across speakers, utterances and contexts.

2.1.6 Exemplar Theory

Arguably, the greatest contribution to the theory of variation up to date has been rendered by Exemplar Theory, which has taken a sensible account of frequency effects on speech perception and production building up a realistic language hierarchy.

Nevertheless except for the basic thesis that speakers operate distinct speech representations of classes, or exemplars, some of the other notions have undergone a considerable evolution over the relatively short existence of this theoretical paradigm in linguistic studies. In particular, one of the main issues under debate is what constitutes a unit of speech perception and production.

In this respect, it is possible to discriminate between "mixed" models, where exemplars and the categories they constitute are viewed as coexistent (e.g. Nosofsky 1986, Pierrehumbert 2001, Bybee 2005); "pure" exemplar models, which only classify separate exemplars (sound types, syllables, pitch accents etc.) according to their distribution (e.g. Walsh et al. 2008) and "extreme" models postulating the absence of any units or types. Under this approach an exemplar is an acoustic sample stored in memory in its global context (Wade et al. 2010).

It seems logical that the last two types of models are only a derivation from the first one. Even if we oust the notions of category, unit and class from the language system, we would still need rules and algorithms for exemplar identification and production. That, in fact, means re-constructing the system.

Apart from that, complete absence of category / unit storage in memory, substituted by acoustic samples, is hardly possible for a variety of reasons. First of all, in second language acquisition basic linguistic knowledge is imposed on a learner. Without it, the learning process would not be at all possible. Even native speakers are confronted with it at school and onwards. This awareness, in its turn, is most probably integrated in the mental representation. People are therefore capable of both subconscious and conscious, i.e. analytical, speech perception and production.

Another important question in exemplar theoretical studies concerns frequent and infrequent classes. The main claim here is that the former are characterized by a large number of exemplars and are most eagerly produced, whereas the latter decay with time if there is no activation (Pierrehumbert 2001; Bybee 2006). It is true that frequency and recency are important factors in speech production. The above studies present comprehensive accounts of these processes. However, frequent categories make up a relatively small percentage in the system. However, getting a full perspective presupposes the knowledge of mechanisms behind the less frequent categories as well.

Noteworthy in this regard are the findings by Goldinger (1998), who pointed out some recurring effects in rare categories, namely that of unusualness. An infrequent exemplar heard in an unusual context and/or produced by an unfamiliar voice is more likely to be stored in memory than otherwise.

Yet, questions remain regarding both classes: why are some very infrequent exemplars still stored in memory and others decay, just as the contexts in which they initially appeared; and why are some frequent instances not stored in memory? The response to the above questions may be that those exemplars are, or are not, behaviourally relevant. Apart from that, speakers vary in individual perception, as well as language ability, which may all serve as a filter and obstacle for exemplar perception and production.

A common assertion about category architecture in exemplar studies is that it is represented by the so-called exemplar "cloud", whose size depends on whether the category is frequent or not. Pierrehumbert (2001) stated that, while categories develop over time, frequent categories accumulate a larger number of exemplars. A similar effect was found by Schweitzer and Möbius (2004) for syllable durations. However, Wade et al. (2010) predicted an opposite trend on the same phenomenon stating that frequent categories are more stable than the infrequent ones and are therefore characterized by a smaller degree of variation. Granted these contradictory findings, it is relevant to find out whether frequency is the key differentiator behind those language events.

If we look at the problem of variation and categorization from a second language acquisition (SLA) perspective, a whole array of additional factors is to be taken into account, e.g. a speaker's internal cognitive processes and strategies, the relative distance between languages and their respective categories, the speakers' previous language experience and so on.

A considerable number of theories and hypotheses emerged, triggered by this complexity of the SLA processes.

2.2 Variation in theories of second language acquisition

Second language acquisition is a complex multidimensional process, which involves both individual variables, treated in detail in Chapter 1, as well as an array of linguistic and extra-linguistic factors. Numerous theories attempted to explain various aspects of this process through *internal factors*, i.e. a speaker's strategies and cognitive processes, and *external factors* – dependence on the source and target language for SLA. We will further call these two types of theories **speaker-dependent** and **language-dependent**. Some theories evolved around the idea of language ability, or rather the inability of L2 speakers to attain a full mastery of a foreign language, which is of special interest for the current study. Such frameworks will be referred to as **ability-dependent** theories.

Not all of these theories deal directly with cross-linguistic intonational variation. However, they do contain important implications for the latter phenomenon. We will therefore present a summary of these approaches as an attempt to answer a few important questions regarding cross-language intonational variation and introduce our own understanding of this aspect of second-language speech[1].

In this respect, the degree of variation appears to be one of the most crucial questions. Is there more variation in an individual's L1 or L2? Is the speech of native or non-native speakers more variable? There can be six logically possible answers to the above questions:

1) an individual's L1 is more variable than his or her L2;
2) an individual's L2 is more variable than his or her L1;
3) an individual's L1 and L2 are equally variable;
4) non-native speech is less variable than native speech;
5) non-native speech is more variable than native speech;
6) non-native and native speech are equally variable.

In other words, there are three possibilities in each case: a smaller, an equal or a greater degree of cross-linguistic within-speaker and cross-speaker variation. We can find conflicting support for all of these options in the three types of SLA theories, i.e. speaker-, language- and ability-dependent.

1 It should be noted in this respect that in the given study we distinguish between the parallel terms relating to this process: learning thus stands for the classroom-based process, whereas acquisition results from a natural exposure in the target-language environment. The terms foreign and second, with regard to language, correspond to the former and latter processes, respectively. We shall make this distinction clear, where necessary. However, the two processes are in many ways similar and often overlap. Therefore, for general descriptions the more widely accepted term second language acquisition (SLA) will be employed.

2.2.1 Speaker dependent theories

2.2.1.1 Basic Variety

One of the most influential theories proposing that second language speech is characterized by a smaller degree of variation on both the intra-speaker and inter-speaker levels is that of Basic Variety, introduced by Klein and Perdue (1997).

In their seminal article the authors pose a fundamental question of whether natural languages could be much simpler than they are. Indeed, there is a considerable amount of structure and logic in any given language, but there is also a lot of, possibly unnecessary, complexity and redundancy. As opposed to this, there exists "*a well-structured, efficient and simple form of language – the Basic Variety"* (Klein & Perdue 1997: 301). The researchers discovered this phenomenon as a result of a longitudinal study involving 40 learners of English, who appeared to develop a relatively stable L2 system to express themselves in a given communicative situation.

The authors further put forth four key assumptions pertaining to the theory of Basic Variety:

A) Language acquisition constitutes a series of transitions from one learner basic variety to another, and this process is highly systematic.
B) A learner variety is characterized by a limited set of organizational principles, which may evolve during acquisition, which, in turn, entails a reorganization of the system.
C) A learner variety is a system in its own right, with a particular inventory of language units and organizational principles. Fully-fledged languages are but borderline cases of the Basic Variety, representing a stable state of language acquisition where individuals stop learning because there is no difference between their variety and the variety of their social environment.
D) Transitions to more complex learner varieties, including the final one, represent an evolutional process and should therefore be studied in a similar way: from the elementary manifestations to the more complex cases.

Following the logic of the theory in question, it appears that second language acquisition is characterized by a gradual increase in the degree of speech variability as the learner's proficiency reaches higher levels. Therefore, learners with a low proficiency level should demonstrate a low degree of variation, whereas the speech of highly proficient individuals should be highly variable approaching the target-language variation levels, i.e. native-speaker performance.

2.2.1.2 Interlanguage

An alternative approach to the theory of cross-language and cross-speaker variation is presented by the Interlanguage Theory, introduced by Larry Selinker (1972).

According to the theory in question, in the course of second language acquisition learners develop intermediate grammars, which are distinct from their native, as well as the target language, and represent independent systems per se. These systems evolve over time as the speakers accumulate experience in processing language input and controlling the respective output as a result of employing various internal strategies.

The above-mentioned notions of independence and evolution are similar to the treatment of L2 speech in the theory of the Basic Variety. However, there is a number of fundamental differences between these two approaches as far as language processing is concerned.

In his model, Selinker introduced several cognitive processes involved in second-language learning:

A) *Language transfer*: some units, concepts and rules are exported by learners from their L1 into the interlanguage
B) *Transfer of training:* some features of the instructional process find their way into their interlanguage
C) *Strategies of second-language learning:* are specific to the learning material
D) *Strategies of second-language communication:* interlanguage units vary specific to the communicative situation and the interlocutor
E) *Overgeneralisation of the target-language linguistic material:* some rules and concepts of the target language are applied onto a large number of phenomena than it is appropriate

Transfer

The notion of language transfer is central to the Theory of Interlanguage and contradicts considerably with the concept of Basic Variety. According to Klein and Perdue, learners develop an effective *error-free* system, which is only simpler in structure than any given fully-fledged language. Yet, a learner system can only be error-free, given that a certain unit is transferred, only if there exists an identical counterpart in the target language, which will be a case of the so-called *positive transfer*. However, if corresponding units in two languages are not identical, this will result in a speech production error, or *negative transfer*. It is indisputable that no two languages are completely identical – therefore errors resulting from transfer should be unavoidable. The concept of overgeneralization also yields an opposite prediction on the phenomenon in question: applying the same rule onto an inappropriate language unit will inevitably result in an error.

Transfer also appears to be crucial for the acquisition of prosody for two main reasons. First of all, by the very nature of suprasegmental events, they can be super-imposed, or mapped, onto any segmental material. It is especially accounted for by the fact that prosodic units share a certain universality (e.g. Bolinger 1978): the falling intonation is predominantly related to the meaning of finality and therefore mainly finalises statements; the rising intonation, on the contrary, is associated with non-finality, so one of its main applications is questions. Secondly, due to the fact that in no natural language are the prosodic units classified into a generally recognized inventory, they are also not taught, or presented, in any way, to the language learners. The latter, in their turn, have to rely solely on their language feeling in order to recognize the structural and semantic aspects of intonation in a given L2.

This feeling, naturally, can often be misguiding: learners may fail to recognize the existing prosodic contrasts in a given target language (Scuffil 1982; Cruz-Ferreira 1989). Such prosodic events will therefore be replaced with the events of the learners' native language (Tarone 1978; Lepetit 1989; Ramirez Verdugo 2003; Mennen 2006), both on the phonological, i.e. as far as the prosodic categories are concerned, and the phonetic level – with regard to the fine-grained inter-unit peculiarities.

Following an extreme formulation, complete transfer of prosodic events into a given target language would mean an equal degree of variation between a speaker's L1 and L2 due to the fact that all the exemplars are actually identical to those in the native language.

However, this case is hardly possible, as it is claimed that other interlanguage processes are at play alongside with transfer. One of those processes, also coined by Selinker (1972), i.e. fossilisation, is claimed to be triggered by transfer.

Fossilisation

Fossilisation is a stage, at which a learner's interlanguage ceases to evolve despite naturalistic and / or classroom-based exposure to the target language. According to Selinker (1972), the main reason for fossilization is transfer. For example, if a given category is transferred from L1 into the interlanguage, it is then employed systematically, and the learner has no need and / or desire to substitute it for an appropriate target-language unit.

It appears to be a highly probable scenario for an L2 speaker to fossilize his or her prosodic idiolect resulting from transfer, especially in the above-mentioned event, when the speakers fail to distinguish between some prosodic contrasts in a given target language. The process in question is certainly easier to observe from a diachronic perspective, but its features can also be seen at a given point through a limited inventory of prosodic events consistently employed by a speaker.

Fossilisation and transfer might also result in the fact that certain target language categories are avoided because they are replaced by L1 units.

Avoidance and overuse

Apart from the above-mentioned reason, avoidance can be explained by the complexity of the respective L2 category. This effect was noticed in a study by Celce-Murcia (1991) who observed a child learning English and French and consistently avoiding the forms, which were physiologically difficult to produce. Still another explanation of avoidance is a speaker's inability to perceive a given L2 category. It appears to be especially true for the prosodic domain where even native speakers have difficulty in discriminating between units and their meanings (e.g. Scuffil 1982); in non-native speakers this results in avoidance (Cruz-Ferreira 1989).

As a compensatory phenomenon, some units may be overused. Support for this phenomenon was found for prosody: in a study by Ramirez Verdugo (2003) Spanish speakers of English consistently overused the falling category and avoided the low rising tone to import the meaning of non-finality.

From a variational perspective, avoidance would certainly entail a smaller degree of variation in a given idiolect. Overuse could also lead to the same outcome if the same category were to be employed in similar realizational variants. However, from an exemplar-theoretic perspective, if an overused category were to become a frequent one, then it would either be stable or variable (see Section 2.4.1 for a discussion).

The short overview of the Interlanguage Theory and its related processes presented above drew attention to a few implications this theory has for the understanding of the phenomenon of variation, in general, and intonational variability, in particular. Nevertheless a separate body of research within the given framework concentrated specifically on the latter issue.

Variation in interlanguage

Selinker et al. (1975) argued that the employment of the cognitive processes and strategies mentioned previously, i.e. transfer, fossilization, avoidance and overuse, results in a certain systematicity in interlanguage, whereby a unique grammar is created. This grammar, or a learner system, is characterized by *systematic variability* (Tarone 1983), which is triggered by context, task or speech style. This argument was supported in several studies of phonological phenomena: in the use of corresponding phonological and syntactic structures (Hyltenstam 1977, 1978; Dickerson & Dickerson 1977 – quoted in Tarone (1983)).

Several studies investigated prosodic phenomena based on these three aspects: the task perspective (Local 1989, 1992; Local, Kelly & Wells 1989; Couper-Kuhlen & Selting 1996; Couper-Kuhlen 2001), the stylistic dimension (Lowry 2002, Grabe 2004) and context (Halliday 1967, Ladd 1980).

The other important type of variation in interlanguage, which has been studied but scarcely this far is that of *nonsystematic variability*. Notable in this regard is the study by Ellis (1985), who investigated the random interchangeability of two negation patterns in his subject's idiolect. The researcher argued that in some situations a learner's grammar was independent of any situational, contextual and stylistic factors. Any given pair, or more, of forms *"will be manifest in the same linguistic environment and they will be used to perform the same communicative function with the same social signification" (p. 53)*. This means that these forms exist in free variation.

This argument was further criticized on the grounds that no two morphological forms can be completely interchangeable (Schachter 1986), which may be true for the grammatical domain, but it appears that free variation is not only possible – it is predominant – in prosody. It is indisputable that there are certain meanings applied to prosodic categories and their free replacement is considerably restricted on the phonological level. However, on the phonetic level, i.e. within a given phonological unit, all that is observed is free variants (bearing in mind the segmental limitations).

This type of free, or non-systematic, variation in prosody still largely remains beyond the focus of attention. Peppe et al. (2000), one of the few studies mentioning the phenomenon in question, distinguished it as one of the several types of prosodic variation – random variability. However, the authors did not elaborate on this subject in the experimental part of the investigation. Nevertheless the importance of examining nonsystematic prosodic variation was acknowledged for the purposes of speech recognition and synthesis (Granstrom 1997), as well as for the diagnosis of aphasic patients (Peppe et al. 2000).

Additionally, it appears that studying the patterning of those free variants would shed light onto the functioning of the higher-order phonological categories in interlanguage.

The Diffusion Model devised by Gatbonton (1978) proposes one of the possible explanations of how interlanguage categories are formed and of their respective content.

2.2.1.3 Diffusion Model

According to the Diffusion Model (Gatbonton 1978), second / foreign language learning is a two-phase process which consists of:

a) *the acquisitional phase* – in which a learner acquires a new form, which coexists with an old one in free variation, the former having first a limited distribution and then spreading over more contexts
b) *the replacement phase* – in which a learner systematically sorts these two forms into their appropriate contexts.

Towell and Hawkins (1993) modified Gatbonton's original model by adding two additional stages to it:

Stage 1: *nonlinguistic strategy* – where a learner does not differentiate between the forms yet
Stage 4: *completion phase* – where a learner has reached complete mastery of a language and everything is "in its place".

Both the original and the modified version account relatively well for the domains they were originally and further applied to: Gatbonton investigated the development of phonological categories; Towell and Hawkins explored the employment of prepositions in French, and Ellis (1985) looked into a particular instance of negation in English. In all these cases, there can be no truly free variation, and the authors, in fact, studied the coexistence and interchangeability of one correct and one incorrect form at a time.

The same parallel can be drawn in the domain of interlanguage prosody where an L1 category (the incorrect variant, in case the two categories are not identical) would coexist with an L2 category in free variation until, if at all, it is replaced by the latter. However, as mentioned earlier, prosody is a unique language subsystem allowing of unconditionally free variation on the phonetic level, of which the Diffusion Model takes no account.

One additional implication can be drawn from the Diffusion Model: the proposed four stages appear to correspond roughly with the three proficiency levels in a learner's development: low – nonlinguistic strategy and acquisition phase; intermediate – replacement phase; and high – completion phase.

It follows logically that the first stage should be characterized by a smaller degree of variability, according to the model, than the other two, due to the fact that a learner's idiolect is limited by a smaller number of exemplars acquired. The variability ratio between the second and the third stage cannot be defined at once. The replacement phase should indeed be characterized by a substantial degree of variability caused by the cumulative coexistence of native and non-native instances, whereas the completion stage would be free of the L1 "noise". Variation on that final stage would depend on the size of the actual phonological inventory and on the density, or non-systematic variability, within each of its constituent categories.

2.2.2 Language-dependent theories

2.2.2.1 Perceptual Assimilation Model (PAM)

The Perceptual Assimilation Model developed by Best (1995) is based on the so-called direct realist view of speech perception. According to this approach, perceptual primitives are the actual articulatory gestures. In the course of the first-language acquisition, an individual gradually gets attuned to constellations of gestural variants in the speech flow. As a result, a speaker learns to identify the higher-order articulatory invariants, which, in their turn, should account for all contextual variations, e.g. speaking rate, cross-speaker differences and allophonic variation[2].

In second language acquisition, *"the fundamental premise of the perceptual assimilation model of cross-language speech perception is that non-native segments … tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space"* (Best 1995: 193). Therefore, typological proximity / distance plays a crucial part in the outcome of the assimilation process, which may follow one of the three scenarios:

1. Assimilated to a native category:
   a. as a good exemplar
   b. as an acceptable but not ideal exemplar
   c. as a notably deviant exemplar

2. Assimilated as uncategorizable speech sound: within the native phonological space, but not as an exemplar of any L1 category

3. Not assimilated to speech (nonspeech sound): heard as a noise or another nonlinguistic signal

Best (1995: 194-195)

Although Best spoke of native and non-native sounds in her model, it appears to be applicable to prosodic events to an equal degree: typologically and phonetically close L2 variants may merge with an L1 category (or will be, in other words, transferred – see Subsection 2.2.1.2), even despite slight or significant dissimilarity; more distant L2 instances may fall beyond any of the

[2] This understanding of invariant, or, in other words, category, formation and architecture is reminiscent of the structural linguistic view on phonological units and their variants.

native prosodic categories – in that case a new category may be formed in the course of SLA; or L2 exemplars will be identified as non-speech (in other words, they will not be perceived, or avoided – see Subsection 2.2.1.2).

The first scenario – the merger of categories – predicts a similar or a greater degree of variation in L2 categories depending on the relative distance between the single tokens: for the former case it should be small or insignificant, whereas for the latter it should be notable. In the second scenario, if a separate category is created for non-native instances, it appears that the degree of variability will depend upon an individual's proficiency level, i.e. the higher it is the more exemplars the category will have acquired – the greater the degree of variability. Finally, the last possibility implies no change of phonetic variation pattern to the source category. However, should a given token represent an L2 category and is perceived as non-speech, this implies a smaller degree of variability on the phonological level due to the non-establishment of an L2 invariant.

2.2.2.2 Second Language Speech Learning Model (SLM)

In his speech learning theory, Flege (1995) dwelled upon the learnability of non-native sounds based on their relative distance to the L1 categories – an approach similar to the Perceptual Assimilation Model discussed above. SLM, however, departed from PAM on a number of important points, as well as introduced several new ideas.

The model is built around four major postulates and seven hypotheses. Of the hypotheses, five contain important implications for the current investigation. We will look at each of them more closely below.

SLM hypotheses:

1. L1 and L2 sounds are related to one another perceptually on the phonetic level, and not on the phonological level.

This idea is different from the one introduced in the Perceptual Assimilation Model, where assimilation is based on matching new events with an L1 invariant. Nevertheless it does go very much in line with the concepts of the Exemplar Theory, viewing categories as dynamic constellations of exemplars.

2. A new phonetic category is established, provided that a learner is capable of discriminating the phonetic differences between L1 and L2 sounds.

The given hypothesis appears to support the findings of several other investigations, quoted previously, where learners failed to discriminate between L1-L2 contrasts unperceivable to them (Scuffil 1982; Cruz-Ferreira 1989).

3. The probability of discerning these phonetic differences increases proportionate to the perceived phonetic dissimilarity between L1 and L2 sounds, i.e. the greater the distance between the sounds in the phonological space – the better the discrimination is expected to be.

This is a new idea compared with the predictions of other theories, especially the Contrastive Analysis Hypothesis (Lado 1957), which states that similar categories are easier to perceive and acquire. SLM thus predicts a difficulty to discern differences between typologically close categories.

4. Equivalence classification is a mechanism blocking category formation, in that L1 and L2 sounds get linked to a single native category and get perceived as equivalent. Such sounds then resemble one another in production.

As a logical development of the previous hypothesis, SLM holds that in a case when a learner fails to discriminate between L1 and L2 contrasts, they merge into a single category. Consequently, native and non-native sounds are produced in a similar way.

5. A non-native learner's categories for a given target language are different from those of a native speaker: they may "deflect" away from the L1 categories in order to maintain contrast between native and non-native events in the common phonological space.

This idea complies with the fourth postulate of the Second Language Speech Learning Model that *bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space"* (Flege 1995: 239). This argument is reminiscent of a study by Flege and Bohn (1989), where non-native speakers tried to demonstrate contrast between L1 and L2 vowels by producing the respective tokens significantly apart from each other in the phonological space.

Summing up, we can draw several implications for the theory of intonational variation from the SLM. First of all, it predicts difficulties for category formation for typologically close languages due to the fact that some L1 and L2 events may be perceived as pertaining to the same native-language category and further produced in a similar way. This idea is close to the notion of transfer developed in the Interlanguage Theory. However, whereas transfer means a complete replication of L1 features in L2 production, SLM only states the similarity of tokens. Moreover, as Flege pointed out, the interference process is bidirectional, so not only do the native sounds affect the production of L2 events, the former are also modified under the influence of the latter.

This implies an increased space for variation within such a merger category, which contains both L1 proper events and a constellation of deviating exemplars.

Secondly, hypothesis 6 implies that non-native speakers will deflect their productions from the target-language ones trying to preserve the contrast between the categories, which, in turn, would mean that a considerable amount of exemplars will be avoided. That would then leave less room for variation in the L2 realizations.

2.2.2.3 Markedness differential hypothesis

Markedness differential hypothesis (Eckman 1977) is a "revised" version of the Contrastive Analysis Hypothesis (Lado 1957), which holds that in SLA similar features are easier to acquire and different elements are, consequently, more difficult to acquire. A number of studies have found counterevidence to this idea (e.g. Moulton 1962). Moreover, other theories, e.g. the Second Language Speech Learning Model (Flege 1995) support just an opposite view on the problem: it is the similar events that present learning difficulties.

Eckman proposed his own explanation for L2 category acquisition, which is based on the Universal Grammar and which should predict the "directionality of difficulty" as well as the "relative degree of difficulty" between a certain pair of phenomena in two languages. In the Markedness Differential Hypothesis this is achieved by introducing the notion of markedness into the model. Thus, the central statement of the hypothesis is that *"a phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A.*" In other words, marked phenomena are "rarer" phenomena in languages when we compare them.

The three postulates of the Markedness Differential Hypothesis read as follows:

1. Those areas of the target language which differ from the native language and are more marked than the native language will be difficult.
2. The relative degree of difficulty of the areas of the target language which are more marked than the native language will correspond to the relative degree of markedness.
3. Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.

(Eckman 1977:329)

At the core, the Markedness Differential Hypothesis is a theory of distributional differences between languages. Therefore, successful second language acquisition and proficiency will follow if a learner understands those differences and is capable of modifying his L1 structure towards the L2 structure, by “redistributing” the features and units. This hypothesis does not give any exact indications as to the degree of variation between an individual’s L1 and L2.

However, the successful learning \ acquisition, and redistribution caused by it, should result in distribution \ frequency patterns similar to those in the target language. Consequently, the higher the proficiency level, the closer the variability patterns should resemble those in the respective L2.

2.2.3 Ability-related theories

2.2.3.1 Fundamental Difference Hypothesis

The Fundamental Difference Hypothesis was introduced by Bley-Vroman (1989), who attempted to answer the question of what the logical problem of foreign language learning is. The author claimed that it is the inaccessibility of Universal Grammar that makes it impossible for adults to master a foreign language as opposed to children, who learn languages naturally and easily, due to their access to the universal Language Acquisition Device (Chomsky 1965), which is not available to adults any more.

One of the author’s major claims is that adults almost inevitably result in failure when it comes to SLA, whereas children are bound to uniform success in this matter. As we demonstrated in Chapter 1, this is but slightly overgeneralized: there are successful adults (hence the array of studies on ultimate attainment in SLA – e.g. Ioup et al. 1994; Moyer 1999, 2004; Abu-Rabia & Kehat 2004; Bongaerts 1999; Bongaerts et al. 2000), as well as unsuccessful children (Flynn & Manuel 1991, Flege et al. 1997, Pallier et al. 1997, Abu-Rabia & Kehat 2004, Moyer 2004, Trofimovich & Baker 2006). According to the author, on the other hand, success in adult learning is somewhat of a pathological phenomenon. This is a commonly held belief; however, the percentage of this „abnormality“ varies from 5 % (Selinker 1972) to 35 % (Seliger 1978), which amounts to a significant part of a certain population of learners.

Another important point set forth by Bley-Vroman is that adults’ learning is characterized by fossilization, whereas there is no fossilization for children. At least one counterargument to this point can be presented by the existence of consistent pronunciation errors in child speech, e.g. their inability to distinguish between „r“, „l“ and „w“ or to pronounce the correct variant of „r“. There is anecdotal evidence that in some cases an individual’s pronunciation may fossilize at this stage of imperfect mastery of a native language, and this person will carry on the lisp throughout his or her life, unless a child receives targeted instruction on this problem and sometimes even despite such external assistance. Rare as these cases may be, they do lie on the other edge of speech „pathology“.

Also related to learning / instruction issue is the idea that children do not require foreign-language instruction and any kind of guided exercise. It might be true that the outcomes of formal instruction vary and that it's role cannot be evaluated definitely. Nevertheless exposure to the target language alone is no panacea for mastering a foreign language as we have shown in Chapter 1. Even in naturalistic acquisition there always exists an element of implicit instruction, through parental or other people's correction and feedback. Moreover, there are many cases where instruction is the only possibility of exposure to a foreign language, as there are many people who get the chance to practice it in the classroom only. The very fact that they manage to acquire a sufficient mastery of a given foreign language, in order to employ it and to satisfy their communicative aims, proves the effectiveness of formal instruction, at least to a certain degree.

It is precisely this type of people, i.e. learning a foreign language in artificial conditions, or with only occasional possibilities of naturalistic exposure, that the Fundamental Difference Hypothesis does not account for. First of all, most people start learning a foreign language long before the end of puberty, and they should therefore still have access to Universal Grammar. However, their exposure to an L2 is limited to several hours of classes per week and maybe interrupted for shorter or longer periods due to some external reasons, during which some of the knowledge and skills get forgotten or deteriorate. An individual may then re-commence learning the same foreign language at a later stage, i.e. at university – as an adult, using general problem solving abilities for SLA instead of the Universal Grammar (Bley-Vroman 1989). In such a heterogeneous L2 exposure / acquisition scenario some people nevertheless manage to acquire a near-native mastery of a foreign language, while others do not. It is important to identify both types of cases, as well as the ratio between them, in order to find out whether success in acquisition is really pathological. One marker of such identification could be the subject matter of the given research – intonational variation – by pinpointing those individuals who actually resemble native speakers in the production of prosody.

2.3 Summary of theoretical approaches to intonational variation

2.3.1 General linguistic theories

In Section 2.1 we explored five major linguistic theories in order to formulate our own idea of intonational variation. Thus, we reviewed the respective concepts within Structural Phonology, Generative Linguistics, Optimality Theory, Prototype Theory and Exemplar Theory.

It appears that Structural Phonology and Exemplar Theory have rendered the greatest contribution to the theory of intonational variation up to date, though with varying implications as to the structure of a given category. The remaining theories under analysis concentrated more on reducing variability or stating its irrelevance, rather than studying it.

Structural Phonology focused on the phonological level of analysis, viewing categories as abstract invariants composed of features and parameters. Variation was viewed as free within a category's boundaries, although any phonetic representations outside those boundaries were not accounted for.

Exemplar Theory, on the other hand, did take account of the phonetic level. More specifically, it postulated that speakers operate speech representations, or exemplars, and not abstract sets of features and parameters. With this approach, even irregular, outlying realizations received explanation.

There are also two important notions for the study of intonational variability in exemplar theoretic studies – those of frequency and infrequency. The main question here is whether frequent or infrequent categories are characterized by a greater degree of variability. One interpretation is that while categories develop over time, frequent categories accumulate a larger number of exemplars and result in a larger exemplar cloud, hence a greater degree of variation (Pierrehumbert 2001, Schweitzer & Möbius 2004). Contrary to this assumption is the idea of relative stability, which is claimed to be characteristic of frequent classes. That, in its turn, would imply a smaller degree of variation within a frequent category (Wade et al. 2010). Even despite these contradictory viewpoints, exemplar theory still provides a realistic explanation of how variation shapes language and speech.

Without a doubt, however, the predictions of both Structural Phonology and Exemplar Theory ought to be tested experimentally.

2.3.2 Predictions of SLA theories about degrees of cross-linguistic intonational variation

For a better understanding of intonational variability in SLA processes, we attempted to summarize the implications on degrees of variation in the speech of non-native individuals, based on the theories of second language acquisition that were reviewed in the current chapter.

Table 2.1 below is meant to reflect the pertinent ideas of all the theories under discussion.

To summarize the predictions of SLA theories on intonational variation, it should be mentioned first of all that the Fundamental Difference Hypothesis was not included in the above summary, as it did not make any predictions on L2 variation: it only postulated an inevitable failure of a non-native speaker facing a foreign language. Therefore, this theory is but a challenge to verify on some actual data and subjects.

In general, however, most SLA theories appeared to predict a smaller degree of variation for low-proficiency speakers (Basic Variety, Interlanguage Theory, Diffusion Model, SLM and PAM). This fact could be explained by the category formation and the corresponding accumulation of new exemplars in an L2 idiolect. A greater degree of variability is expected if L1 and L2 categories are mixed in an idiolect due to their typological proximity and the resulting non-discrimination by speakers (Interlanguage Theory, Diffusion Model, PAM, SLM). This pattern is expected at higher-proficiency levels, when a substantial amount of L2 exemplars have been processed. The ratio between the variability levels at the so-called stages of average (intermediate) and above-average (advanced) proficiency is not completely clear. It is almost definite that the higher the proficiency, the closer a speaker is to the native realizations and variability amounts.

Thus, Basic Variety, Diffusion Model and Markedness Differential hypothesis would all predict the highest degree of variation for the above-average individuals: the first theory – due to the expansion of variety; the second one – because of the sorting out of the L1 „noise“and further „pure“ category enrichment; and the final paradigm – at the completion of the learning process for the markedness relations. However, at some average stage there might be a greater degree of variability simply caused by the L1 / L2 language mixture, in which case above-average proficiency speakers might have a richer variety in their realizations, if they manage to master a greater amount of L2 exemplars.

Theory	**Predictions on degrees of variation in L2 speech**		
	greater	smaller	equal
Basic Variety	for higher-proficiency speakers	for lower-proficiency speakers	---------
Interlanguage	for higher-proficiency speakers: in case of partial transfer, when L1 and L2 units are mixed	for lower-proficiency speakers: categories may be mixed due to transfer, but there are fewer accumulated L1 exemplars than in more proficient speech	in case of complete transfer of identical categories from L1 into L2
Diffusion model	for higher-proficiency speakers	for lower-proficiency speakers	---------
Perceptual Assimilation Model	for higher-proficiency speakers: at a stage when an L2 sound is assimilated to an L1 category; and the greater the distance between tokens, the greater the variability	for lower-proficiency speakers: at the stage of establishment of a new category, when exemplars are being accumulated	---------
Speech Learning Model	for higher-proficiency speakers: if an L2 sound is assimilated to an L1 category	for a) lower- and b) higher-proficiency speakers: a) during the category development stage b) after the establishment of a category due to deflection effect - some borderline tokens are avoided in production	---------
Markedness Differential Hypothesis	---------	---------	for more proficient speakers: managing to reach a native variability level

Table 2.1: Predictions on variation degree between native and second-language speech based on the SLA theories reviewed in Chapter 2

2.3.3 Conclusion

The analysis of general linguistic theories on intonational variation in the current chapter has established the relative contributions of different theoretical paradigms to the phenomenon at issue. Thus, we concluded that Exemplar Theory and Structural Phonology have provided two valid approaches to the description of intonational categories: the former focusing on the phonetic level; the latter – on phonology.

We also reviewed the major SLA theories and their treatment of variation, especially with regard to proficiency, which, as we have stated in Chapter 1, is most closely correlated with pronunciation aptitude – the core notion in the current investigation. The main implications of those frameworks were then summarized in Section 2.3.2.

All of the above theoretical considerations have enabled us to propose our own idea of how the phenomena of foreign language pronunciation aptitude and proficiency relate to the degrees of intonational variation. In Chapter 3 we will proceed with presenting the corresponding hypotheses.

Chapter 3

General hypotheses on degrees of intonational variation in native and non-native speech with varying levels of pronunciation aptitude and proficiency

The title of the current chapter is not particularly succinct. In fact, it can't be, because we are attempting to approach a number of phenomena from different disciplines: intonational variation in its relation to foreign language pronunciation aptitude, and, further on, the relation of pronunciation aptitude to proficiency. In what immediately follows, we will propose our solution as to the interaction between these domains.

In Chapter 1 we reviewed the phenomenon of pronunciation aptitude, as it is reflected by the modern studies, and have arrived at the conclusion that this kind of talent cannot be equaled to or encumbered by any factors, especially external. Thus, the amount of classroom or naturalistic exposure cannot guarantee a success in an L2. We have reviewed examples when, even given a substantial amount of foreign language experience, the speakers did not make a sufficient progress in a foreign language (Flynn & Manuel 1991; Flege et al. 1997; Pallier et al. 1997; Abu-Rabia & Kehat 2004; Moyer 2004; Trofimovich 2005), and, vice versa, they were able to achieve high proficiency even despite the less intense or even sporadic foreign language input (Ioup et al. 1994; Moyer 1999; Bongaerts 1999; Bongaerts et al. 2000; Abu-Rabia & Kehat 2004). The internal factors, such as the degree of motivation as well as the speaker's age were not reliable predictors of the success in foreign language pronunciation either (Colletta, Clement & Edwards 1983; Clement & Crudenier 1985; Flege 1987; Piske et al. 2001). Rather, pronunciation aptitude always remained to be intact and set apart from all the other factors.

However inconsistent the effects of external and internal variables of the L2 performance may be, language aptitude can only be observed most objectively when all the other factors are leveled to the maximum degree. Following this logic, the influence of language aptitude would be most evident in homogeneous groups, where the age, motivation and amount of exposure are almost equal. In those cases, it would only be the amount of natural ability that would enable the speakers to reach higher proficiency levels.

The latter, in their turn, appear to correlate with the degree of language variability, as suggested by the summary of second language acquisition models introduced in Chapter 2.

Thus, low variability would correlate with low proficiency, according to the majority of theories, due to the fact that fewer tokens have been mastered by the speakers at early stages of acquisition. Later proficiency levels are characterized by a mixture of L1 and L2 exemplars, with the decaying amount of native language tokens, rising levels of proficiency and the corresponding accumulation of L1 proper tokens. Thus, a highly advanced speaker should possess a better command of an L2, enabling him or her to employ a greater variety of exemplars than an average speaker. For the latter, in turn, language mixing takes place: an average learner's idiolect is populated by some native-language exemplars, alongside with some L2 proper events. The more an individual acquires of L2-specific tokens, the more of them should be used. In an opposite case, if a speaker is not able to distinguish L1 / L2 contrasts, then the native categories should just be exported into the foreign language.

Thus, we have established a relationship between proficiency and the degree of L2 intonational variability. The remaining link to establish is that between proficiency and pronunciation aptitude.

For that purpose, we should assume a homogeneous sample of speakers, with relatively equal external and internal variables, exposure being of special importance hereby. The degree of proficiency, acquired by any single one of such speakers, should then actually demonstrate his or her natural ability for learning the phonetics of a foreign language, as compared to other individuals in this group. Further on, various degrees of proficiency across the speakers would correlate with various degrees of language, and intonational, variation, as follows from the conclusions of Chapter 2 and arguments provided above. Finally, we can logically associate the variation patterns established for different proficiency levels with the corresponding classification of pronunciation aptitude.

We would consequently assume a low level of pronunciation aptitude for low-proficiency individuals; below-average aptitude level for speakers of below-average proficiency; an average aptitude would be defined for average-proficiency individuals; above-average proficiency would stand for above-average aptitude; and, finally, highly-proficient individuals would be described as being highly capable in their L2 pronunciation aptitude.[3]

Based on the considerations laid down in the current chapter we would like to propose our hypotheses as to the degrees of intonational variation in the realizations of speakers of varying pronunciation aptitude.

Hypothesis 1: L2 speech of less proficient (lower than average level of pronunciation aptitude) speakers should exhibit a smaller degree of variability in a foreign language.

3 We ought to stress that the postulated hypotheses only apply to the specified homogeneous samples of speakers, having received equals amounts of L2 exposure and being relatively equal on the motivation scale.

Hypothesis 2: In an opposite fashion, advanced-proficiency (higher than average level of pronunciation aptitude) individuals should be more variable in their speech due to their higher acuity in mastering L2 exemplars and employing them in communication.

Hypothesis 3: Average-proficiency (and average-aptitude) speakers are expected to exhibit variation levels between the more and less talented individuals.

We would also like to propose a hypothesis, competitive to the third one, due to the fact that some theories imply an uncertain relationship between the variability degrees on the average level of pronunciation aptitude and above.

Hypothesis 4: Degrees of variation between of average and higher than average pronunciation aptitude are unpredictable and can only be verified through the analysis of real-life speech data.

Our final hypothesis should account for the predictions of some models, i.e. Basic Variety and Markedness Differential Hypothesis, that the final proficiency destination for non-native individuals is a mastery level approaching native-speaker performance.

Hypothesis 5: On a proximity scale, speakers of higher than average aptitude and proficiency should approach most closely the native-speakers' variation pattern. Average as well as lower than average levels of aptitude and proficiency should correlate with lower degrees of variability, establishing a greater distance of such speakers from the proper L2 performance.

In the current and previous chapters we presented our general ideas, as well as hypotheses, on the degrees of variation in L2 speech pertaining to speakers of various degrees of proficiency and aptitude. In order to carry out actual data analysis, it appears to be indispensable to select a suitable instrument, namely, an intonational paradigm, which best suits a cross-language investigation of intonational variation of this kind.

For this purpose, in Chapter 4 a review of the major theories in the description of prosody will be presented, in view of proposing such a model or, alternatively, a combination of models.

Chapter 4

Prosodic models

A wealth of comprehensive studies have approached the phenomenon of speech prosody, in order to provide an explanation of its functioning. This effort has resulted in a variety of models, clustering around a number of conceptual paradigms, which all resolve some issues in the description of suprasegmentals. However, these frameworks differ both in their methodological and theoretical standpoints, as well as predominantly focus on one particular language at a time.

For the current investigation, however, it is important to adopt a paradigm, or several paradigms, best suited for cross-linguistic analysis of intonational variation. For this purpose, we will review the existing prosodic models in an attempt to select such a paradigm.

Classification of models

The existing intonation models have been traditionally classified into theoretical (perceptual – Pike 1945; impressionistic, proto-phonological – Ladd 1998) and experimental ones (instrumental – Pike 1945; instrumental, phonetic – Ladd 1998), based on the main instrument, or foundation, on which they have been developed – either the analysis of earlier research and the subsequent proposal of a new conceptual paradigm or the investigation of speech corpora for testing and advancing a set of ideas. In the latter case, this may, in a similar fashion as for the theoretical models, result in a new phonetic-phonological framework. Alternatively, some researchers avoid the phonological level in their analysis, only concentrating on the acoustic properties of the speech signal without any meta-specification. Thus, we assume that modern intonation models, for the mere sake of structured analysis, may in fact be subdivided into three groups: theoretical, experimental-phonological and experimental-phonetic.

4.1 Theoretical models

The prevalent theoretical approach to the description of intonation was developed actively since approximately 1940-s - 1950-s of the last century – earlier studies contain but sporadic comments on the phenomenon at issue – within the general structural linguistic paradigm. It is still widely used for applied, i.e. teaching, purposes. A distinction is traditionally made between the British and American schools of the given approach, due to some important conceptual differences.

4.1.1 British and American descriptions in comparison

One of the most crucial differences between the two schools was first formulated by Bolinger (1951) as a dichotomy of *levels vs. configurations,* or *static* vs. *dynamic* elements (Lieberman & Pierrehumbert 1984) and became the subject of debate for intonation researchers in the decades following.

Within the American school of research, intonation was predominantly described in terms of discrete pitch levels, and not only by the structural linguists. Representatives of other paradigms also adhered to this understanding, which we will demonstrate further in the chapter. Pike (1945) proposed one of the most comprehensive accounts of American English intonation system postulating four significant levels, labeled 1 through 4 in the transcriptions, which specified *extra-high*, *high*, *mid* and *low* levels, respectively[4]. According to the author, these levels were not associated with any absolute measures; instead, the significance of any particular pitch-level could be revealed only through its relative height in comparison with the other pitch-levels. In particular, Pike stated that *"the distance between the four levels is not mathematically fixed, uniform, or predictable. It varies from individual to individual, and the individual varies his own intervals from time to time."* (Pike 1945: 26). Moreover, it was posited that pitch levels have no meaning of their own, due to the fact that the intonation contour as a whole is a structural and semantic unity. The pitch levels thereby mark the beginning points of the *F0*[5] direction change in a given contour, thus being its building blocks. Falling and rising tones consist of two contour points, according Pike: one delineating the beginning and the other – the end of the contour. Three-point tones would also be possible, in cases where the voice goes down and then up again on the same phonetic word[6]. In general, Pike classified all American English contours as falling or rising from level *x* to level *y* within the four-level spectrum introduced above.

Bolinger (1958), as well as several other scholars (e.g. Crystal 1969, Pierrehumbert 1980, Lieberman & Pierrehumbert 1984), criticized this approach, first of all, due to the lack of experimental evidence that the four proposed levels actually represent phonological contrasts. Secondly, the postulated relativity of pitch-levels is in itself a questionable concept, as it is not clear to which degree it can expand if an individual varies pitch from utterance to utterance. It is therefore a matter of debate whether the levels are mutually exclusive or overlapping, and whether in the latter case it is possible to distinguish between them.

[4] Trager and Smith (1951) also supported the idea of a four-level analysis of American English intonation. Additionally, they put forth four additional variations within each of these levels.
[5] *F0* stands for fundamental frequency – one of the basic physical properties of the speech signal.
[6] With a phonetic word we mean a single word or a group of words clustered around a single accent and tone associated with it.

The intonation research within the British school, on the other hand, developed around the notion of the *F0* dynamics, i.e. rises, falls and plateaus in pitch contours. It appears that the two approaches to the description of intonation are in some ways two sides of the same coin: it is indisputable that a pitch contour is a sequence of upward, downward and level configurations (absolutely any visual representation of *F0* will demonstrate that), but those configurations pertain to certain levels within a given pitch range. A valid description of intonation would then define phonologically relevant units within the prosodic system of a given language, as well as the possible degree of phonetic variation within and between those established categories.

Another cornerstone difference between the "classical" as well as later frameworks in the British and American traditions concerned the basic unit of intonation. For both schools it was the pitch contour, but its composition was different. The British school investigated "global" contours, i.e. units spreading over longer parts of sentences or whole phrases; whereas American scholars predominantly analyzed intonation in terms of sequences of "local" contours associated with separate stressed syllables. In Pike's model sentences were also split into smaller contours. However, not all stressed syllables in a phrase shared the same relative weight: the last accent was considered to be the most prominent one, semantically and phonetically. That idea was equally present in the British English descriptions, where the most important and the only compulsory element in a contour was the nucleus, i.e. the last fully stressed syllable[7].

Phonologically, any contour, be it global or local, was described as having several structural components, which were similar in the British school and Pike's description of American English. As mentioned above, the focal centre was occupied by the *nucleus* (or the stressed syllable of the last contour in a phrase in Pike's model), which was also characterized by the most prominent pitch change in a given contour, realized through one of the available tones in the inventory of a given language.

For British English, different authors proposed slightly varying tonal inventories with the agreement as to the four basic tones: *fall*, *rise*, *rise-fall* and *fall-rise* (O'Connor & Arnold 1961; Halliday 1967; Crystal 1969; Kingdon 1973; Sweet 1878), with the exception of Palmer (1922), who did not include rise-fall in his description. Most scholars introduced some notion of pitch-level into the tonal classification, i.e. they discriminated between *high* and *low* falls and rises (O'Connor & Arnold 1961; Sweet 1878), and Crystal (1969) also recognized the *medium* pitch-level, as well as brought out the notion of range – *wide*, *normal* and *narrow*. Alongside with bidirectional tones (fall-rise and rise-fall), some authors set forth tridirectional pitch movements: *rise-fall-rise* (Crystal 1969; Kingdon 1973) and *fall-rise-fall* (Crystal 1969). Apart from that, all the complex tones, i.e. bi- or tridirectional, were described as falling into two varieties: *undivided* and *divided*, differing by their expansion – on one or more than one word, respectively (O'Connor & Arnold 1961; Halliday 1967; Crystal 1969; Kingdon 1973). Finally, the *level* tone

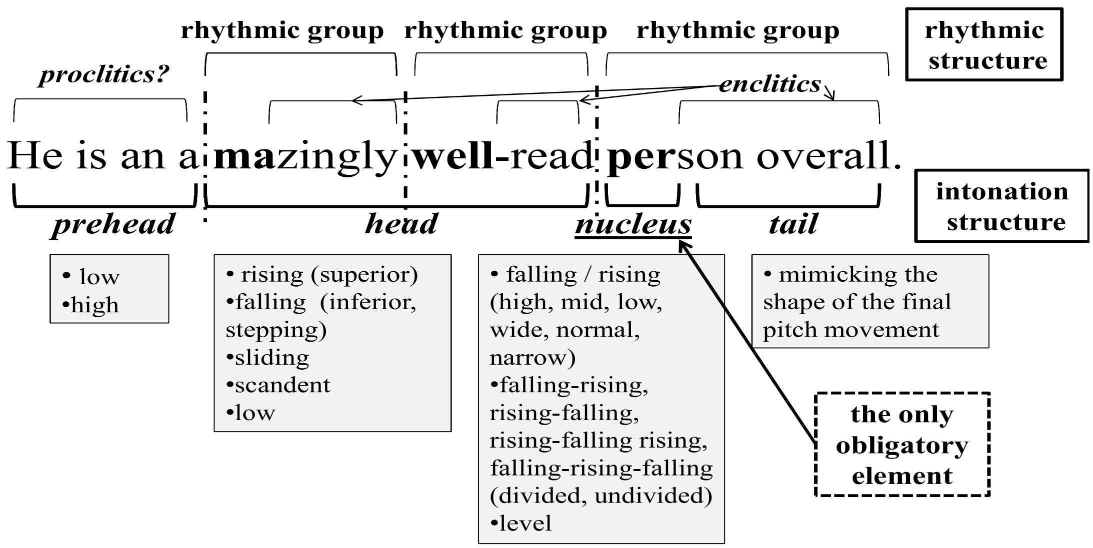
[7] Not all American school approaches recognize the special status of the nuclear intonation events. For a detailed discussion refer to Section 4.3.1 below.

was singled out by Sweet (1878) and Crystal (1969) as a special type of plateau-like final pitch movement. On top of this relatively complex classification, Kingdon (1973) divided all tones into *normal* and *emphatic*.

The syllables following the nucleus would form the *tail* of an utterance, and they would normally carry on the pitch configuration, set out by the nucleus: the *F0* would continue to rise after a rising-type nucleus, whereas it would fall after a final falling pitch element – part of a uni-, bi- or tridirectional tone. The nucleus and the following tail were treated separately from the rest of the stressed and unstressed syllables in a phrase, if there were any. Those, in their turn, would make up the *head* of a tune, except for the syllables preceding the first stressed syllable (or the *onset* - Crystal 1969), commonly referred to as the *prehead*, which could be either *high* or *low* depending on whether a speaker should commence an utterance at the top or bottom of his or her pitch range. The latter variant was said to be the default for English. The stressed syllables making up the head of an utterance were claimed to be pronounced on a level pitch (high or low) and they therefore carried the *static* tones, according to Kingdon (1973), as opposed to the *kinetic*, or dynamic, tones specific to the nucleus. The level and arrangement of stressed syllables in a head would determine its type in the general classification as *rising* (Crystal 1969; superior – Palmer 1922), where the syllables would form an ascending pattern, or *falling* (Crystal 1969; inferior – Palmer 1922; stepping – O'Connor & Arnold 1961), where there were a step down from one stressed syllable to another and which was claimed to be the norm for English. In addition to that, some authors distinguished *sliding* (O'Connor & Arnold 1961) and *scandent* heads (Palmer 1922): here, each stressed syllable would be characterized by a falling or rising pitch movement, respectively. Crystal (1969), however, pointed out that this may just as well be successions of falling or rising nuclei.

Further on, any head would be split into *rhythmic groups* (rhythmic units – Pike 1945), which consisted of a stressed syllable and the unstressed syllables preceding it, or *proclitics*, and following it, or *enclitics* (pre- and postcontour in Pike's classification, respectively), if there were any. These two notions are still a matter of debate, as it is often difficult to say, in a case when there is more than one rhythmic group in an utterance, whether a given unstressed syllable belongs to the previous rhythmic group as an enclitic or if it is part of the next group being its proclitic. Some scholars proposed a semantical-grammatical solution to this dilemma (Karnevskaya 2007): groupings should be based on the lexical and morphosyntactic proximity of a given unstressed syllable to the stressed one.

British Model



Pike's Model

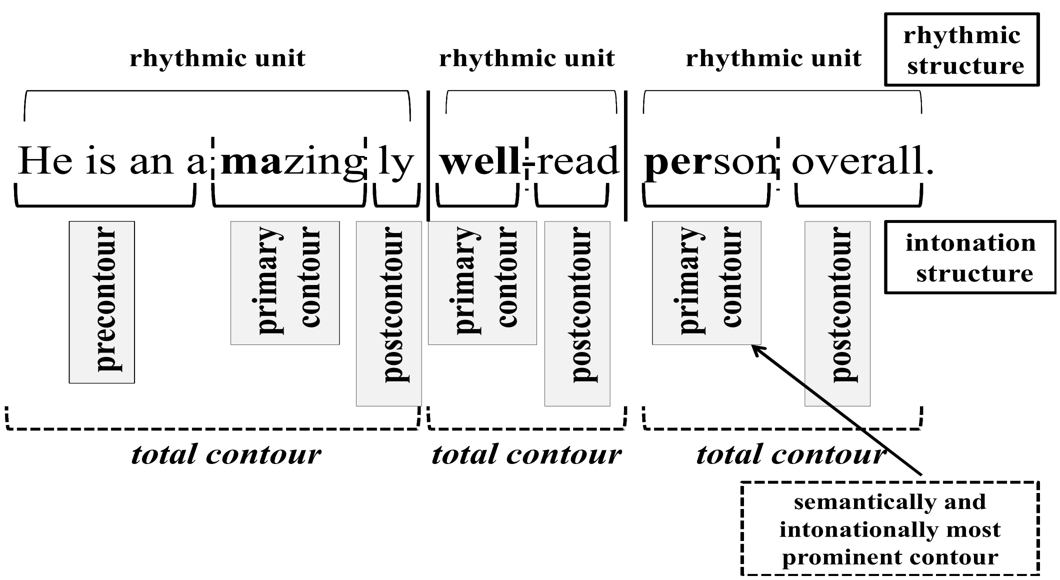


Figure 4.1: Schematic representations of intonation units in the British school and Pike's (1945) models

However, if proclitics and enclitics are to be prosodic concepts, it is only the suprasegmental parameters that ought to determine the status of these syllables. The other approach questions the very existence of proclitics for English and other Germanic languages (e.g. Kohler 1977, Möbius 1995) and postulates the so-called left-headed rhythmic pattern, where a rhythmic group only comprises the stressed syllable and enclitics, regardless of any lexical or grammatical ties between the words.

Opposed to this pattern is the right-headed rhythm, common for Romance languages, where only proclitics are said to exist. One may question the existence of proclitics in the head for all Germanic languages, but enclitics alone cannot account for the syllables in the prehead, which should either be treated separately altogether or the notion of proclitics ought to be preserved. In any case, this question is beyond the scope of the present investigation.

Still another questionable idea regarding English rhythm in the structural phonetic descriptions was the hypothesis that all rhythmic groups should be relatively *isochronous*, i.e. pronounced within equal time intervals, which was criticized in experimental studies that did not support this trend (e.g. Shen & Peterson 1962; Pierrehumbert 1980).

The prehead, the head, the nucleus and the tail taken as a whole represented an *intonation unit* in the British English tradition, called a breath group by Lieberman (1967), a tone-unit by Crystal (1969) and a sense group by O'Connor & Arnold (1971). The differing terms also reflected the deviating approaches to the definition of this unit: biologically – as a speech abstract pronounced on a single breath; phonologically – through the defined tones; or semantically – as a coherent chunk of text. Minimally, an utterance would consist of one intonation unit, but, in fact, there could be multiple intonation units in a phrase. In Pike's description of American English there was no notion of the head, or the intonation unit. Instead, he analyzed separate local *total contours* (comprising a precontour, a primary contour and a postcontour) granting a special role to the final contours in an utterance, like the British scholars, which presented a slightly different approach to the intonation unit structure, but, in principle, shared several ideas proper to the British school.

Thus, the commonalities between the classical British and American descriptions included the following: the special status granted to the last prominent syllable in a phrase and the general structure of the rhythmic group. The tonal inventories also appeared to be similar if one translated Pike's notation (for a detailed account refer to his classification) into British school terms. In fact, the idea of tonal configurations was not totally absent from his description as he used the terms "rise", "fall" and "fall-rise" in a similar way it was done by the British scholars. In turn, the concept of pitch-levels found its place in the latter framework as well, where up to three pitch-levels were postulated for the tones.

4.1.3 Kohler's (1977 & 1995) models of German intonation

An analogous intonation description, as the one proposed by the British school, was further developed for German by Kohler (1977). In his model, the author put forward a hierarchical structure of prosody, where strong and weak syllables were to form the basic building blocks. Those syllables constituted *tacts* (cf. rhythmic groups in the British tradition), consisting minimally of one strong syllable, though, more frequently, of a strong syllable and a number of enclitics. Thus, tacts in Kohler's model were "left-headed" and did not contain any proclitics. The next level in the prosodic hierarchy was the *intonation unit*, which comprised one or more tacts. Similar to the British school description, Kohler discriminated between the nuclear and prenuclear part of the intonation unit. The latter was not specified in much detail in the model. However, the author listed five possibilities for the realization of the prenucleus in German: as a low or high level sequence of syllables; as a gradually ascending tune; and, finally, as two "jumpy" melodies – successions of falling or rising pitch movements. The nucleus of the intonation unit would contain the nuclear syllable and the tail, if any. The final, and most important, pitch movement on the nucleus could be classified within the six major tonal types:

1) the falling tone – characterized as reaching the bottom of the pitch range
2) the high rising tone – ascending to the top portion of the pitch range
3) the low rising tone – going up to the middle of the pitch range
4) the mid level tone – realized as unchanging pitch
5) the falling-rising tone – a bidirectional configuration reaching the middle of the pitch range on its rising element
6) the rising-falling tone – a bidirectional configuration ending at the bottom of the pitch range

Kohler also described some of the typical contexts, in which the above tones could be used. Apart from that, he claimed that not all combinations of the prenuclear and nuclear parts are possible: for example, the rising tones were only supposed to have a level, high or low, or a gradually descending prenucleus.

As seen from the summary above, Kohler's model closely approached the British school description, both conceptually and with regard to the terminology.

Möbius (1985) carried out a number of experiments in order to verify whether German intonation could indeed be defined within Kohler's model. The results of the experiments were not straightforward. On the one hand, the subjects were able to identify the postulated tone inventory. However, no instances of the mid level tone were found in the experimental corpus, which implies that it cannot be stated with certainty whether Kohler's inventory accounts for the whole variety of intonational categories in German. Möbius also tested the notion of the nucleus in the German intonation. Like the other experiment, that investigation did not yield any conclusive outcome: there was no absolute consensus between the subjects as to the position of the nuclear syllable.

Based on this theoretical account of German intonation, Kohler (1995) further developed an applied model for text-to-speech synthesis, enriching the paradigm with some new features.

Thus, in general, sentence stress attained a four-level gradation into reinforced, neutral, partially and completely deaccented. Intonation categories were described as a dichotomy of pitch peaks and valleys, where the former would contain either a unidirectional fall or a fall-rise and the latter would form a low or a high rise.

The updated model also postulated some rules of pitch category concatenation in connected speech, i.e. the principles of how peaks and valleys are combined with regard to the stressed and unstressed syllables, on which they are realized. Additionally, the paradigm specified the alignment of peaks and valleys onto syllables – early, medial and late.

In contrast to the original description, Kohler added the category of the prehead, defined as the syllables preceding the first sentence stress, into the system with the possible values of “high” and “low”.

With regard to the global intonation trends, the author adopted the concept of downstep, which determines the structural pitch lowering from one sentence stress to another, regardless of the time that elapses between them. Kohler denied the idea of declination for German intonation, adopted in other models, for example, the IPO approach, which will be presented in Section 4.2.4.1. Finally, the model accounted for several types of prosodic boundaries between the intonation units, as well as included adjustments for different degrees of speech rate.

In conclusion, it should be noted that the two theoretical models of British and American prosody, as well as Kohler’s initial description of German, appear to be well-suited for their initial purpose – teaching suprasegmentals to non-native speakers, as these approaches provide sufficient, if not superfluous, amount of detail about the prosodic system. One can, in fact, develop a phonologically varied intonational idiolect in an L2, based on these paradigms.

On the other hand, British and American approaches were often criticized as being impressionistic and not linked to any actual data. More specifically, as mentioned earlier in the chapter, the pitch-level approach was judged as failing to provide a phonological distinction between the categories, and it was therefore doubtful whether the four postulated categories are discrete. Nevertheless pitch-level models were developed in later American studies, where these shortcomings appeared to have been overcome. The British school descriptions, in their turn, were questioned for the number and nature of tones and tunes in the description, as experimental studies did not find any data representing some of these categories. Kohler’s second model, however, proves that it is possible to apply British-school models into experimental research and speech technology.

4.2 Experimental-phonological models

Experimental-phonological models combined theoretical depth with sufficient data verification. We will therefore present two types of such models in the current chapter: autosegmental-metrical approaches and *F0* approximation models.

4.2.1 Autosegmental-metrical approaches

Ever since it was first introduced by Goldsmith (1976, 1990), autosegmental-metrical phonology has been one of the central approaches in theoretical descriptions. The kea idea behind this paradigm is the proposition that phonological representations are not linear: they consist of several tiers composed of strings of segments[8], each representing one feature, e.g. segmental, tonal, rhythmic, etc. Those features, in their turn, as well as their constituent segments, were linked by certain association rules. For instance, a syllable segment on the syllable tier would be associated with a given tonal segment on the tonal tier.

Autosegmental-metrical phonology further developed into a framework of models, including a substantial number of prosodic descriptions for various languages. By far, one of the most influential phonological models has been the approach proposed by Janet Pierrehumbert (1980) for American English intonation.

4.2.1.1 Pierrehumbert's phonetics and phonology of American English intonation

Pierrehumbert constructed her model by incorporating the ideas of Liberman (1967) and Bruce (1977), as well as several structural phonetic concepts of Pike's (1945) model and some ideas of the British school. The model has undergone several developmental stages: from the original system laid down in Pierrehumbert's dissertation (1980), to a slightly revised version for applied synthesis purposes (Beckman & Pierrehumbert 1986; Hirschberg & Pierrehumbert 1986), and, finally, to the ToBI framework (Silverman et al. 1992; Pitrelli et al. 1994). The first two stages were conceptually very close, whereas for the ToBI (Tones and Break Indices) the model was revised considerably. It is this final paradigm that was further applied to a substantial number of languages, including German. We will discuss the ToBI convention separately in due detail, as the current investigation has adopted a modified ToBI approach to intonation labeling.

[8] The notion of segment in Goldsmith's formulation indicated a single unit on one of the tiers (e.g. H or L for high or low tones on the tonal tier; a phoneme on the segmental tier; or a syllable on the syllable tier), and not a sound in the dichotomy of segmental versus suprasegmental features.

First, however, we would like to introduce the original model, its modifications, as well as the subsequent criticisms and alterations by other scholars.

In her dissertation, Pierrehumbert postulated a phonological system consisting of three major components. The first one was the metrical structure of a given text, which indicated stressed and unstressed syllables specifying the relative strength of these syllables. It should be noted that Pierrehumbert did not introduce any prominence gradation, other than the dichotomy between the strong (stressed[9]) and weak (unstressed) syllables. She did mention, on the other hand, that it is possible to derive a finer-grained classification of prominence from the metrical structure. All the strong syllables were treated equally, including the last stressed syllable. This contradicted the idea of the nucleus, as the last prominent syllable and the only obligatory element in an intonation group, in the structural British school description. For Pierrehumbert, the only difference of the last strong syllable, as contrasted to the other ones, would be its mere position, and not any other attributes.

The grammar of phrasal tunes, consisting of sequences of low (L) and high (H) tonal targets, was the second element in the model. Both strong and weak syllables would carry a tonal target, whereby the former would be marked with a star (*) and the latter – with an upper hyphen (¯). All the strong, i.e. starred, syllables were marked with pitch accents, of which there were seven in Pierrehumbert's original notation: H* – a simple high target on the accented syllable; L* – a low-pitched strong syllable; L*+H¯ – a low target on the accented syllable followed by an immediate rise to a high level on the syllables following; L¯+H* – a high-pitched stressed syllable immediately preceded by a rise from a low level; H*+L¯ – a high-level target followed by a low trailing tone; H¯+L* – a step down onto the accented syllable from a high-pitched preceding part and H*+H¯ – a high-pitched target on the stressed syllable which spreads onto the immediately following unstressed syllables.

Tonal targets were not only associated with stressed and unstressed syllables; additionally, they would mark phrasal boundaries, i.e. points at the end of an utterance or within an utterance before a pause. These events were treated independently from the pitch accents, as it was claimed that the pitch movement at the end of a phrase was not determined by the shape of the pitch accent: rather, its configuration would align with the phrase edge – forming boundary tones in the intonation system. It should be noted in this respect that the pitch movement at the very beginning of a phrase would also be marked with a boundary tone indicating the level, at which an utterance starts. Per default, as it was posited, a phrase would begin with a low initial boundary (normally not marked in the notation). The reverse utterance start, at high pitch, would also be possible, reminiscing of the notion of the high prehead in the British school notation, which comparison Pierrehumbert also acknowledged herself. An additional element would

[9] The notion of stress in Pierrehumbert's model pertained to sentence and text level, i.e. it was the relative prominence that determined whether a given syllable would be classified as strong or weak. Consequently, word stress as such would be just a potential for prominence.

account for the complexity of pitch configuration before the final boundary tone – the phrasal tone. The latter would align approximately with the last syllable of the word carrying the nuclear accent, or its position might be delayed, if there would be multiple syllables before the boundary. Such a definition implied that the phrasal tone might be placed anywhere between the last syllable of the nuclear word and the boundary tone, which did not yield any exact location for this element of the intonation structure. In the revised model (Hirschberg & Pierrehumbert 1986), the phrasal tone acquired a new definition as the part of the phrase spreading between the last pitch accent and the boundary. Still, the exact alignment of this tone remained unclear. Another peculiarity pertaining to the phrasal tone was the fact that it was subject to spreading, in which cases it determined the *F0* configuration before the boundary – as either unchanging, plateau pitch or as a change in pitch direction after the phrasal tone.

The boundary tone would be placed just before the rightmost, as well as the leftmost, boundary. Both the phrasal and the boundary tone would only be unidirectional – high or low. In special cases, where the postnuclear part would be considerably long and would in fact constitute a tag of some sort, multiple phrasal and boundary tones would be possible in the model.

The final third element of the phonological representation were the rules for aligning the available tunes with the text. Those were context-sensitive local rules, applied left to right, specifying the possible modifications to the *F0* contour.

The general intonation structure of discourse, according to Pierrehumbert, incorporated the three above elements, i.e. the metrical structure of the text indicating strong and weak syllables, the tonal targets applied to those syllables, as well as the alignment rules, to form well-shaped *tunes* for each particular utterance. Any tune, in its turn, would consist of one or more *intonation phrases* (IP's) – the basic macro-units of intonation description. The structure of an IP would incorporate one or more metrically strong syllables carrying pitch accents (the focus of the model), as a rule, one phrase accent following the last pitch accent and one (or rarely more than one) final, as well as optionally an initial, boundary tone. In a revision of the original model (Beckman & Pierrehumbert 1986), a new macro-unit was introduced and, namely, the *intermediate phrase* (ip) – a coherent chunk of text not followed by a full-fledged boundary. The phrasal tone would then apply to ip edges, whereas the boundary tone would only pertain to the IP level. By creating this hierarchical structure of an intonation phrase containing one or more intermediate phrases delineated by phrasal tones, the authors resolved the issue of phrasal tone alignment: temporally, both the latter and the boundary tone would coincide; structurally, however, they would belong to different levels.

In general, the grammar of the model determined the sequences of pitch accents, phrase and boundary tones, which were considered legal (see Figure 4.2 below).

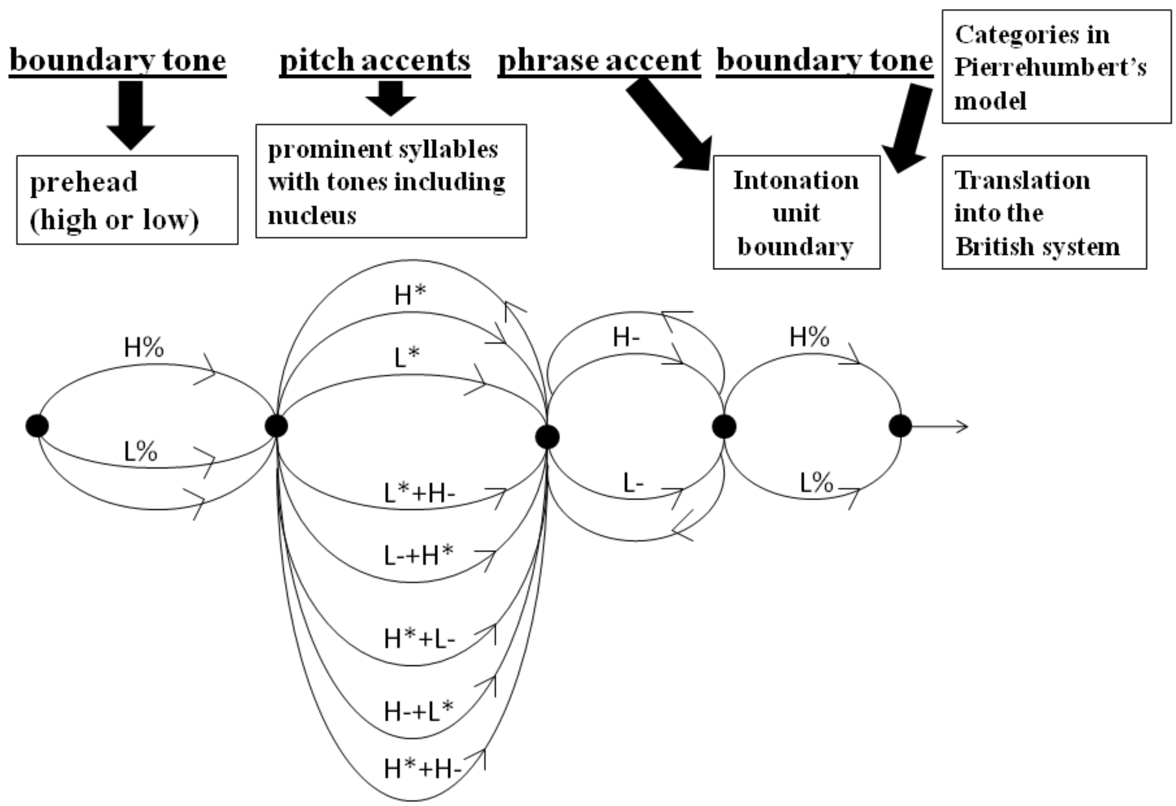


Figure 4.2: Schematic representation of the finite-state grammar for the generation of intonation phrases (IP) in Pierrehumbert's (1980) model. The counterpart concepts between this model and the British school descriptions are connected by arrows

As evident from the above description and the finite-grammar scheme, Pierrehumbert's model indeed included several ideas proper to the classical structural models. Thus, the composition of an IP concurred with the structure of an intonation unit in the British school description, and the notion of the initial boundary agreed with the definition of the prehead. The disagreement between the British and Pierrehumbert's paradigm lay in the treatment of the final boundaries: in the former model the postnuclear part pitch movement would be dependent on the nucleus; in the latter account, however, this configuration was postulated as independent from pitch accents, which presented a more flexible explanation of complex pre-boundary movements common for the English intonation. Similarities with earlier American school approaches are also quite evident: the idea of strong syllables carrying pitch accents and unstressed syllables – either leading or trailing tones appeared to agree with the concept of total contours in Pike's (1945) model. Nevertheless unlike the classical American school descriptions, Pierrehumbert's model reduced the number of pitch-levels to two minimally distinct ones on the phonological level – high and low, as the author pointed out that most English phrases would not contain instances of all four tones in four-tone theory.

Another area of interest in the model were the properties of the *F0* contour between the metrically strong syllables. Pierrehumbert treated those as phonetic, rather than phonological, rules regulating connections between the adjacent pitch accents. In that respect, there could be dipping, spreading and ascending configurations for the transitions, although the author did not provide any detailed description of these mechanisms.

One of the most important ideas in Pierrehumbert's model was the fact that there was no global layer of phrasal intonation. Instead, as mentioned above, all the alignment rules were local and applicable from left to right, i.e. local events would determine the overall trends. In particular, only the *F0* level of the first pitch accent could be subject to free choice; all the subsequent pitch accents would be scaled relative to that first strong syllable, taking prominence relations into account. In addition, all bitonal accents were defined as triggering downstep, a clear *F0* level lowering, of the following pitch accents.

Another crucial element in the phonological representation defining the overall configuration of an intonation phrase was the baseline. The results of Pierrehumbert's experiments implied that the lowest *F0* values are relatively constant for an individual's and that they decline toward the end of an utterance. Therefore, the corresponding values of the pitch accents in an IP were computed relative to this descending baseline. Consequently, although global tendencies were ruled out in the definition of the model, they were still part of it, which is accounted for by the concept of downstep and declination, operating independent of the local events.

Apart from that, pitch range employed for rendering various kinds of pragmatic connotations was still another global phenomenon capable of modifying the values of local events.

Pierrehumbert presented the first comprehensive autosegmental account of American intonation, inspiring later investigations and evoking some comments, developments and criticisms to the model.

4.2.1.2 Other autosegmental-metrical approaches to the description of English intonation: Ladd (1998) and Gussenhoven (1983)

Ladd

Ladd (1998) presented a critical account of Pierrehumbert's model with a revision of several ideas, the main one being the status of the nuclear pitch accents. The author argued that it is not only the mere position of the nucleus that makes it prominent, although in many cases it is indeed located close to a boundary. More important than its placement, according to Ladd, is the fact that the nucleus is the focal point in an intonation phrase. Therefore, prenuclear and nuclear pitch accents may not be put on the same structural scale.

In Ladd's account, the intonation of tags, i.e. relatively long parts of text after the last nuclear accent and before the boundary, also acquired an alternative explanation. Pierrehumbert would permit multiple phrasal and boundary tones to mark such cases. Ladd, on the other hand, would differentiate between postnuclear parts with and without prominent syllables in them. Should any of the words in that position contain strong syllables, Ladd would treat them as carriers of a separate type of pitch accent – the postnuclear accent. By adopting such an approach, Ladd fell back on the results of several other experimental investigations (e.g. Grice 1995; Vella 1995), reporting on secondary prominent syllables after the nucleus. He also quoted a phenomenon, commonly known in the British school tradition as the fall-rise divided (O'Connor & Arnold 1961; Halliday 1967; Crystal 1969; Kingdon 1973), where a falling nucleus is separated from a final rise by multiple intervening non-prominent syllables. The fall was thereby described as the major element, whereas the rise would be somehow subordinate to it. In Ladd's opinion, the latter element could be described as having a postnuclear accent. Worthy of mentioning in this respect is also the notion of partial stress by British scholars. Kingdon (1973) in his division of all tones into kinetic and static ones mentioned that of the latter a high and a low variety would be possible, and they would acquire the status of a full and a partial stress, respectively. Low partial stress would then take any position in an intonation phrase: in the prehead, in the body, as well as in the tail after the nucleus, if prominence is to be attached to this or that syllable. The postnuclear accent is therefore, though a debatable phenomenon, but still one discussed in various studies across linguistic paradigms. On the contrary, if the postnuclear part does not contain any prominent syllables, then there would only be a phrasal and a boundary tone associated with it, according to Ladd.

To summarize Ladd's idea of an intonation phrase structure, it should be noted that he viewed it as a compound phenomenon, containing elements of an essentially different nature: prenuclear, nuclear and postnuclear.

The prenuclear material, as well as the potential pitch accents associated with it, would be a matter of a free choice, only dependent on the metrical structure of a given text. The postnuclear part, on the other hand, albeit containing postnuclear accents, would be closely connected to the nucleus – the central element in the intonation structure.

Figure 4.3 demonstrates Ladd's revisited finite-state grammar to incorporate the compound intonation phrase structure he proposed. Noteworthy is the fact that this grammar does not contain the H*+H¯ accent present in Pierrehumbert's original model, because it was eliminated in the subsequent descriptions (e.g. Beckman & Pierrehumbert 1986).

A third final criticism of Pierrehumbert's model presented by Ladd was the definition of the downstep. According to Pierrehumbert's revised model (Beckman & Pierrehumbert 1986), all bitonal accents were treated as downstep triggers. Ladd argued, however, that the phenomenon in question cannot be treated as only being caused by positional and structural factors. In his account, Ladd gave examples of identical utterances only different by the presence and absence of downstep in them. Ladd's major contention was that "*downstep is an independent intonational choice*". Therefore, the author proposed that the phenomenon at issue be treated as a separate feature, existing parallel to the type of accent per se. He therefore proposed to use a separate diacritic to single out downstepped accents from normal ones.

In general, as evident from the discussion above, Ladd's fundamental work provided alternative explanations to several disputable phenomena in Pierrehumbert's model.

Gussenhoven

Gussenhoven's (1983) model, on the other hand, focused on several prosodic phenomena not touched upon by Pierrehumbert in her studies.

The first one of these phenomena was the *F0* peak timing. In particular, Gussenhoven supported the opinion that the alignment of pitch peaks accounted for the category identification. Should the local maximum be associated with the prominent (starred in Pierrehumbert's notation) syllable itself, then the pitch movement should be perceived as being of a falling variety; if, on the other hand, the target should be shifted, it should create the effect of a rising-falling configuration. The latter, in its turn, may vary as to the exact location of the delay. Further on, the later the delay in the syllable, the more meaningful it would make the respective word. Gussenhoven related this claim to the early structural linguistic definition of the rise-fall-rise as the intensified version of the rise-fall, both structurally, as the peak was more delayed in the former than in the latter, and semantically – the connotation attributed to the rise-fall-rise was in a way a combination of the meanings of the rising-falling and the falling-rising tones (e.g. Crystal 1969).

Gussenhoven's other area of interest was tone stretching – lengthening of the mapped syllables, or stylization, mostly common for the so-called calling contour, which, does not fall to a low pitch-level, but rather stays in the mid range. Therefore, Gussenhoven, along with Ladd (1998), considered it a stylized variant of the fall, equivalent to the notion of the level tone in the British school tradition (Crystal 1969; O'Connor & Arnold 1973). The fact that this type of the falling nuclear movement ends somewhere in-between the high (H) and low (L) targets, Gussenhoven proposed that in such cases the two tones should merge into a third category – mid (M). Pierrehumbert's approach would treat such cases as an H-L boundary.

Gussenhoven also proposed a theory of transitions between the metrically strong syllables – an area only scarcely touched upon in the previous autosegmental studies. Gussenhoven explained the different bridging of pitch accents by the degrees of completion of the respective tonal targets. Thus, the author classified all tones into the following three categories:

- completed tones – all the level targets realized;

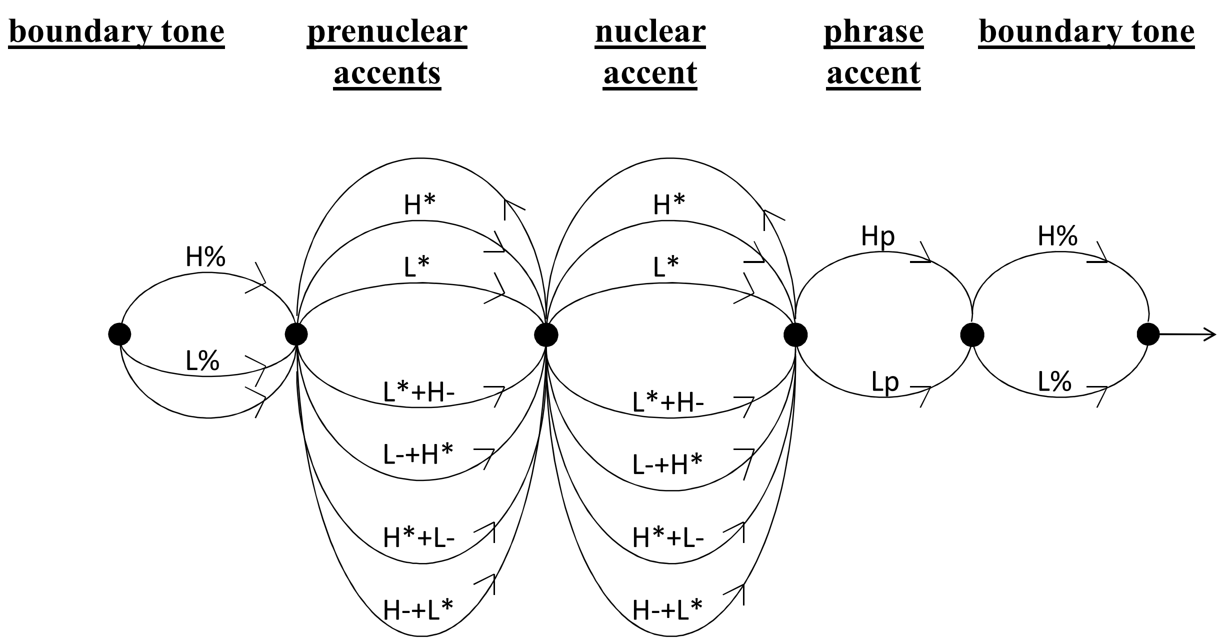


Figure 4.3: Schematic representation of Ladd's (1998) reviewed finite-state grammar of the original Pierrehumbert (1980) model

- half-completed tones – the final element of the first tone merged with the first element of the next tone: e.g. in the case of a combination between H*LH / H*L the final H in the first element would be merged with the initial element of the falling accent;
- completely linked tones - complete elimination of the trailing tone, e.g. in a sequence of H*L / H*L, where the final L of the first element would be removed to merge with the following pitch accent.

It appears that the idea of tone linking in Gussenhoven's approach is closest to the idea of the head in the classical British school tradition, where the prenuclear syllables were treated separate of the nucleus and in general received secondary attention. It logically follows that they could be fully or partially realized in speech. Kingdon's (1973) idea of static tones appears to approach this idea most closely, as it implies the importance of the pitch level of the prominent syllable, regardless of the unstressed syllables. Gussenhoven actually analyzed several types of heads from the British school description, i.e. rising, high, falling and low, in his theory of pitch accent transitions. However, all the pitch accents under investigation were of the two- or three-target type, monotonal categories were not mentioned in the description at all. It appears, though, that if one views all pitch accents as potentially bitonal, one eliminates the truly plateau-type "flat" peaks, where the *F0* remains on the same level between the prominent syllables, e.g. in a sequence of two simple high targets. Those might not necessarily be underdeveloped bitonal or tritonal pitch accents, but intended peaks of stable pitch. Apart from that, some inter-category distinctions might become unclear in the course of linking: in the above example of half-completed tones (H*LH / H*L) it would be difficult to say, whether the first element is a truncated fall-rise or a completed fall. Rather, there would be a way to explain this theoretically, but it could be a challenge to establish a difference between the two on an actual data set.

Finally, Gussenhoven was interested in the phenomenon of the pitch range, especially with regard to the *F0* span of various tonal categories. According to the author, they could spread over a low or a wide range, and undergo or fail to undergo shrinking. There are frequent references to this phenomenon in the British school descriptions, with the references to high and low varieties of tones (see Section 4.2.2 for a detailed account). Crystal (1969) gave the most extensive classification in this respect, where all tones were subdivided into high, mid and low with a further gradation into wide and narrow varieties.

To summarize the autosegmental-metrical approaches to the description of English intonation, it should be noted that they inspired applications of this theoretical paradigm onto a substantial number of other languages. Several of these studies, in particular, approach the German prosodic domain.

4.2.1.3 Autosegmental-metrical models of German intonation: Féry, Wunderlich and Uhmann

Féry

One of the most comprehensive autosegmental-metrical accounts of German intonation was proposed by Caroline Féry (1988, 1993).

When describing German intonation, she departed from the corresponding descriptions of English in several respects. In particular, she posited that there is no phrase accent in German, as normally the final pitch movement is an abrupt fall or a rise, which mimics the configuration of the preceding pitch accent. Consequently, only the boundary tone is necessary for the description of German, according to the author. Even the low boundary tone could be eliminated, if the postnuclear pitch movement were to follow that of the previous pitch accent. There might be, however, a facultative high boundary tone. English, on the other hand, requires both categories, while it possesses a richer inventory of postnuclear realizations than the German language, as Féry explained.

Like Ladd (1998), Féry made a distinction between the prenuclear and the nuclear pitch accents. She pointed out two peculiarities of the nuclear events as opposed to the prenuclear part: first of all, nuclear tones display more variety in their realization, and, secondly, they can be at least bitonal. Prenuclear tones, on the other hand, are either simple-target, or the movement associated with the trail tone in them may be reduced.

Féry recognized six types of nuclear accents:

1) H*L, the simple falling movement, was one of the most common patterns, typical of declarative sentences. Phonetically, two realizations would be possible: if the fall is relatively early in an IP, it should be rather sharp; if, on the other hand, it is relatively late in the IP, it should be realized as a smooth fall or rise.
2) L*H, or the simple rising tone, expressing incompleteness in both questions and declarative sentences. In a sequence of two rising tones, the second one can undergo upstep, i.e. start on a higher level than than first rise.
3) H*LH%, the falling-rising tone, where the presence of a high rising boundary instead of a simple trailing tone is explained by the fact that the final rise always occurs on the last stressable syllable in an IP.
4) L*HL, the rise-fall, or delayed peak in Ladd's (1978) description, expressing some sort of assertion.
5) H*M, the stylized contour, where the high-level target of the nuclear syllable is spread up to the penultimate syllable and all the syllables are slightly lengthened.[10]

[10] Féry adopted Gussenhoven's approach to the description of stylized contours.

6) HH*L, a high tone associated with a metrically weak syllable immediately preceding the metrically strong syllable.

Like Gussenhoven (1983), Féry recognized the phenomenon of linking for the prenuclear pitch accents in German. In particular, the author claimed that the prenuclear monotonal pitch accents (H* and L*) can be thought of as being derived from underlyingly bitonal nuclear tones which are prevented from running their full course, because they are linked to the following nuclear tone. Thus, the linking can be partial, in which case the trail tone of the first nuclear tone would associate with a syllable adjacent to the second one; the unaccented syllables between the two tones would then be unassociated; so that interpolation can take over. The linking can also be complete: the trail tone of the first bitonal nuclear tone would then be deleted and the interpolation would occur between the pitch accent of the first former nuclear tone and the second pitch accent.

The linking phenomenon was also present in Féry's definition of hat patterns, of which she claimed there were two types in German. The first one contained a high prenuclear pitch accent, derived from an underlying H*L, and a proper H*L pitch accent. The author noted with respect to that hat pattern that it were somewhat unstable, i.e. it could be split into two parts depending on the degree of emphasis attributed to the first accent, the speech rate, etc. The second hat pattern, on the other hand, appeared to be more connected, according to Féry, because it consisted of a rising pitch accent L*H bringing about its semantics of incompleteness combined, as in the first hat pattern, with a falling tone H*L.

Féry adopted a similar approach to downstep, as the one introduced by Pierrehumbert (1980). In particular, she posited that, structurally, downstep is triggered by bitonal pitch accents. Semantically, there are two realizational possibilities: if two accents are realized on the same pitch-level, or if the first one is slightly higher in *F0*, the second one should be perceived as more prominent; if, on the other hand, the second pitch accent is downstepped, then both of them are equal in prominence, each forming an intermediate phrase. Therefore, unlike Pierrehumbert's (1980) original model, but similar to Ladd (1998), the author posited that downstep is optional.

Wunderlich

Wunderlich (1988) focused on the interrelation between prosody, syntax, and the information structure of discourse, especially sentence modality and focus-background relations. The author posited a six-tone inventory of intonation patterns for German. The only monotonal accent in the system was the simple high target H*, or the default accent, in Wunderlich's definition. There was also one bitonal plateau pitch accent H*H, claimed to be common for enumeration. Finally, the paradigm contained a bidirectional bridge accent H*HL*, employed for contrast. The rest of the intonational types were postulated in combination with a boundary tone, with which they would typically occur. A special sequence of a high initial boundary and a low target tone %HL* was singled out for exclamations. The two other combinations of a final boundary tone and a pitch accent were proper to the expression of non-finality: L*H% - for a low accented target rising to a high pitch-level toward the phrase break and L*H(H%) – for a steeper rising tone, common for echo questions.

Apart from describing the inventory of pitch accents coupled with boundary tones, Wunderlich also introduced the idea of markedness to all the intonational categories. The unmarked realizations of boundary and trailing tones would then be of a low-target variety, whereas the default for a pitch accent is a high starred tone. Vice versa, utterances ending in a high boundary tone, as well as high-level transitions between prominent syllables and low-level pitch accents would be marked, thus, less common in the German intonation system.

Uhmann

Similar to Wunderlich's model, Uhmann (1988, 1991) concentrated on the information structure of discourse, especially the focus domain. However, unlike any other autosegmental-metrical accounts discussed previously, the author posited the default initial boundary on a mid pitch-level, with two optional alternatives in the low (L%[11]) and high (H%) portions of a speaker's pitch-range. The final boundary tone, how or low, was obligatory in the system, as a phonological correlate of semantic and syntactic phrasing; the initial boundary was an optional element in the paradigm. The inventory of pitch accents consisted of four categories: L* - simple low tone, H* - simple high tonal target, L*+H – the rising accent, H*+L – the simple fall.

Like Féry (1988, 1993), Uhmann recognized bitonality as the only possible realization of the nuclear accents, whereas prenuclear accents could be uni- or bidirectional and would only include trailing tones. The leading tones were not necessary for the German intonation, according to the author.

Uhmann also correlated the postulated intonation categories with the components of the information structure. Thus, as mentioned above, boundary tones would indicate discourse chunking. Pitch accents, in their turn, would reveal focus and topic elements: H*+L would the only definite marker of the former, whereas H* could indicate both; L*+H in the prenuclear position and L* pointed out topic constituents.

Uhmann also mentioned the typical combinations of pitch accents and boundary tones with regard to the types of communicative contexts, in which they are employed: H*+L L% - for

[11] In Uhmann's original description the low tone (L) is called Tiefton (T), but we preserved the common notation for the sake of clarity.

declaratives and wh-questions; L*+H H% - for echo questions and yes/no-questions; H*+L H% - for yes/no-questions; L*+H L% - for wh-questions.

4.2.1.4 The ToBI annotation framework

Autosegmental-metrical phonology inspired the creation of a prosodic annotation convention, ToBI, which was largely based on Pierrehumbert's model (Pierrehumbert 1980; Beckman & Pierrehumbert 1986) for American English.

In line with the general principles of autosegmental-metrical phonology, a ToBI representation for any particular utterance was split between four tiers (Beckman & Hirschberg 1993): orthography, tones, break indices and miscellaneous indices. The orthographic tier contained words and their right-edge boundaries, aligned with the corresponding speech sample. The break indices were to be inserted after each word contained on the orthographic tier and would correspond to the relative degree of juncture between these words. On the miscellaneous tier a transcriber would mark all cases of discontinuity in a signal (noise, coughing, laughing, breathing etc.).

Of all the tiers in a ToBI representation, the tonal tier was the central element, since it rendered the actual information about the intonation events. These events could be of two types in a ToBI transcription: *phrasal tones* associated with intonational boundaries and *pitch accents* – tonal events related to the prominent syllables. Both pitch accents and phrasal tones might only have two values: L (low) – for targets located relatively low within a given pitch range; and H (high) – assigned to the events found towards the higher limit of the pitch range.

On the phrasal level, ToBI presupposed a hierarchical two-component structure constituted by intermediate and intonational phrases. Intermediate phrases were claimed to be more closely connected with each other than intonational phrases, which were supposed to be more independent units. At minimum, any given sentence should contain one intonational phrase composed of one or more intermediate phrases. Each of the intermediate phrases would be delimited from the next one by an intermediate boundary tone, or phrasal tone. A full boundary tone would be found at the end of an intonational phrase. Since the boundary of the last intermediate phrase in a given intonational phrase coincides with the latter, both a phrasal and a full boundary tone should be added at the end of an intonational phrase.

With respect to pitch levels, both types of phrasal tones could take one of the two specified values – high (H) or low (L). The default for the initial point in a given intonational / intermediate phrase, or the initial boundary tone, was mid to low pitch range, and it was not labeled in ToBI transcriptions. Only the less frequent realization of the phrasal onset on a high pitch level (%H) would be noted. Full final and intermediate boundary tones could also be realized either high or low within a given pitch range. In ToBI notation, the pitch level indication

(H or L) would be followed by a minus sign for intermediate phrases (L-, H-) and by a percentage sign for full final boundary tones (H%, L%). As mentioned above, there had to be in fact two tones at the end of an intonational phrase, for it is simultaneously the final point in an intermediate phrase. In that regard, ToBI would allow all possible combinations of pitch levels in intermediate and full final boundary tones (L-L%, L-H%, H-H%, H-L%).

The inventory of pitch accents in American English contained five categories, according to ToBI. There were two monotonal accents associated with prominent syllables: H* - a high or mid pitch tonal target (corresponding with H* and H*+L in Pierrehumbert's (1980) model), L* - a low-pitched tone. The remaining ToBI pitch accents were bitonal. L*+H stood for a low tone target on the prominent syllable, which is immediately followed by a rise. L+H*, on the contrary, signified for a sharp rise onto the accented syllable for a low pitch level. Finally, H+H* was a step down onto the accented syllable from a higher pitch.

The original American English ToBI conventions thus translated the ideas of autosegmental-metrical phonology into a framework for transcribing intonation, which was further applied to other languages. A recent overview of ToBI systems is found in Jun (2005) and contains accounts of Greek, Dutch, Serbo-Croatian, Japanese, Korean, Mandarin, Cantonese, Chickasaw, Bininj Gun-Wok, Italian, Swedish and German.

A ToBI based convention for British English, the IViE labeling Guide, was developed as part of an annotation effort for a corpus comprising nine urban dialects of English spoken on the British Isles (Grabe 2001).

There were several major differences between this British version of ToBI and the original American English paradigm. First of all, an additional phonetic tier was introduced into prosodic transcriptions, which captured the pitch levels of the prominent syllable, the syllable immediately preceding it and all the syllables before the next prominent syllable. Each of those syllables would carry a high (h), low (l) or mid (m) pitch level target. This annotation tier cannot be regarded as purely phonetic, since it still presupposes a substantial level of categorization.

The phonological tier, similarly to the original American English conventions, was represented by pitch accents composed of high and low pitch targets. The inventory of pitch accents was, however, not identical to that of American English. Thus, in addition to the H*, L* and L*+H accents present in the original ToBI, there was also a falling H*L (with a downstepped variant !H*L), a falling-rising L*HL and a rising-falling H*LH tone.

Regarding the phrasal tones, only full intonational boundary tones were part of the system. Intermediate phrases and boundary tones were not recognized as units. Instead, the target tier was supposed to carry all the necessary information on pitch changes within and between prominent syllables. With respect to the set of full boundary tones within the IViE, it included the high and low initial boundaries as well as the so-called "default" boundary tone, signifying no distinct pitch movement after the last pitch accent.

For German, two ToBI models were independently devised – the Saarbrücken and the Stuttgart systems, following different theoretical premises.

The Saarbrücken model (Grice & Benzmüller 1994) very closely followed the original model of American intonation (Beckman & Hirschberg 1993) described above, i.e. the inventories of pitch accents and boundary tones were identical.

The Stuttgart model (Mayer 1995), on the other hand, integrated the autosegmental-metrical intonational analysis of German intonation done by Féry (1988, 1993) described in Section 4.2.1.3. Thus, the Stuttgart model departed from both the Saarbrücken and the American English fameworks to a substantial degree. Most pitch accents in the model were bitonal: H*L – fall (and its downstepped variant !H*L), L*H – rise; HH*L – early peak; L*HL – rise-fall, late peak; and H*M – stylized contour. The monotonal pitch accents, H* and L*, were claimed to stem from the so-called linking processes – the idea was adopted by Féry (1988, 1993) from Gussenhoven (1983), as stipulated in Sections 4.2.1.2-3. It was assumed that all pitch accents are inherently bitonal, but they can completely lose or lend their trailing tone to the next prominent syllable in connected speech.

The treatment of boundary tones in the Stuttgart framework also substantially differed from the original American English model. Following Féry's supposition as to the absence of phrasal accents in German, due to the fact that the final pitch movement mimics the configuration of the preceding pitch accent, Stuttgart ToBI postulated only one possible label for an intermediate boundary tone, marked by a single minus sign. A low (L%), high (H%) or default (%) final boundary tone would be added at the right edge of an intonational phrase. The final boundary type, or the default boundary, signified instances of absence of any pitch change after the previous pitch accent, the same way as it was specified in the IViE convention.

Despite the dissimilarities between the Stuttgart and Saarbrücken models, the authors of these approaches devised a unified GToBI, German Tones and Break Indices, model (Reyelt et al. 1996; Grice et al. 2005). As far as pitch accents were concerned, the unified model almost replicated the Saarbrücken inventory, with the exception of the H+L* accent, a step down from high to low pitch of the prominent syllable, which was added to the system. There were a number of changes in the definition of boundary tones. Only the high (H-) and low (L-) intermediate tone categories were carried over from the Saarbrücken model, although the new model also introduced the downstepped variant of the high intermediate boundary tone (!H-), signifying a step down as opposed to the pitch level of the preceding pitch accent. The full low final boundary tone (L%) was eliminated from the system altogether and replaced by L-% - a low plateau stretch, which may be followed by a drop to even lower pitch. There was also a counterpart high plateau boundary tone H-%. The remaining two boundary tone categories described the final rising pitch configuration: from high to even higher pitch - H-^H%; and from low to mid pitch level – L-H%.

Summarizing the autosegmental-metrical phonological approaches, we ought to point out that the given paradigm has provided both the theoretical basis and an instrumental framework, ToBI, for the exploration of inter-language intonational variation. The INTSINT system constituted an alternative description of prosodic events in different languages with the aim of creating a universal prosodic alphabet.

4.2.4 Hirst and di Cristo's exploratory paradigm for the transcription of intonation: INTSINT

The International Phonetic Alphabet proposed by the International Phonetic Association (the IPA) is an acknowledged convention for transcribing sounds in all languages of the world. An equivalent for annotating intonation is by far non-existent. The ToBI convention, based on Pierrehumbert's model (Pierrehumbert 1980; Beckman & Pierrehumbert 1986), described in Subsection 4.2.1.4, is one of the attempts to propose such a prosodic alphabet.

The INTSINT (International Transcription System for Intonation) model is still another attempt to provide a paradigm for describing intonation events in all world's languages. However, unlike the ToBI framework, INTSINT does not assume a preexisting inventory of intonational categories. Rather, it is an exploratory framework for gathering data on underdescribed languages (Hirst & Di Cristo 1999; Hirst et al. 2000).

In several ways, INTSINT is similar to ToBI. In particular, the main elements in the model are pitch targets (cf. the notion of pitch accents in autosegmental-metrical approaches), which are linked by a transition function. The values of adjacent pitch targets are defined locally, i.e. in relation to the previous accent, similar to Pierrehumbert's model (1980). In that respect, they can be defined as being on the same level, higher, lower, downstepped or upstepped, as opposed to the left-context prominent syllable. The distinction between a higher or lower versus upstepped or downstepped accent, accordingly, would be the pitch change value: the *F0* excursion would be steeper for the former transitions than for the latter.

The top and bottom components of the model would delineate the extreme points in a speaker's voice range. In general, a coherent sequence of pitch targets and their transitions, optionally separated by a pause from the rest of the discourse context, was defined as an intonation unit – an equivalent notion to the one postulated by the British school and similar to the notion of an intonation (intermediate) phrase in autosegmental models.

INTSINT posited a somewhat different definition of initial and final boundaries for an intonation unit. Thus, the default initial boundary would pertain to the mid level in a speaker's pitch range, whereas the *top* (high) or *bottom* (low) values were considered to be the marked realizations of the category in question.[12] The final boundary could have five basic values of *top*, *bottom*, *higher*, *lower*, *upstepped*, *downstepped* or the unmarked variant – the *same*.

The authors pointed out that their, rather detailed, universal description of intonation events may not apply fully to any given language, and it would be necessary to single out the relevant features in each individual case. In Hirst and Di Cristo's survey of twenty languages (1999), more than half of the descriptions were made using the INTSINT notation.

4.2.4 *F0* approximation approaches

A set of intonation models was based on the approximation of *F0* contours in order to derive a set of pitch prototypes for a particular language. These prototypes would then serve as a base for the generation of a certain random pitch contour, pertaining to a given class, but differing slightly parametrically from the prototype. Such models concentrated mostly around the well-known IPO intonation approach.

4.2.4.1 The original IPO model

The IPO (Instituut voor Perceptie Onderzoek) model ('t Hart & Cohen 1973; 't Hart & Collier 1975; 't Hart et al. 1990) was initially developed as a perceptual approach to intonation, to be further applicable to speech technology, and, more specifically, text-to-speech synthesis.

The authors' underlying proposition was that listeners are only sensitive to a highly restricted class of relevant *F0* changes – the so-called macrointonation. Those changes would be the pitch configurations that listeners / speakers intentionally produce. The remaining *F0* information was considered to be "irrelevant detail", or microintonation, according to the authors. Thus, t'Hart and colleagues proposed a simplification, in other words, stylization, of any particular *F0* contour by means of omission of microintonational effects with the aim to extract an inventory of discrete pitch events. The interpolation between the relevant points in an *F0* contour was a straight-line approximation, applied on a declining baseline.

The original IPO intonation model, as well as all the other similar paradigms developed further, followed a strict procedure of extracting the minimally sufficient inventory of relevant prototypical contours. First of all, a given corpus of recorded speech was resynthesized. Afterwards, each resynthesized original was stylized in such a way as to contain the smallest

[12] Compare, for example, the autosegmental models of English and German intonation, which we presented in Section 4.2.1, as well as the classical British and American approaches, where the phrase-initial default was low-pitched. Perhaps the reason for a mid unmarked initial boundary in INTSINT is the fact that it is not bound to any particular language.

number of straight-line segments to represent the relevant *F0* movements, which make the new version perceptually equal to the original. As a result, a set of close copies was extracted. Whether the latter were indeed indistinguishable from the originals was verified by a set of independent listeners. Thus, if a close copy were judged to be a successful imitation of the original, the sample was retained in the database. The verification process further allowed to single out the relevant pitch contours in the corpus, defined as coherent strings of pitch movements, of which there could be several in a given utterance.

‘t Hart and colleagues further proposed a number of dimensions, along which those pitch movements can be analyzed: direction (rising or falling), timing within a syllable (early, late, very late), rate of change (fast, slow) and size (full, half). The specific content and character of these dimensions can be different for various languages due to some universal constraints, e.g. the scholar experimentally defined four possible, and perceptually relevant, pitch range gradations.

Though pitch was implicitly present in the model, IPO scholars did not recognize the phonological relevance of pitch levels per se. ‘t Hart and co-authors argued that if *F0* levels were the generic targets of intonational production, then the pitch movements connecting those levels would have invariant melodic properties, i.e. a fixed rate and timing. This is, nonetheless, not the case in speech. Therefore, pitch movements have some properties, which may not be easily predicted from the interpolation between two given pitch-level targets. The authors further argued that even if pitch level should be included in the typology of the relevant pitch movements, it cannot substitute them as a phonological unit.

Finally, the authors created a grammar of all allowed combinations of the relevant pitch movements. Unlike Pierrehumbert (1980), ‘t Hart and coworkers contended that this grammar is highly restrictive, i.e. by far only a limited number of combinations is allowed for a particular language, thus creating a finite set of common *F0* configurations. The latter were further subdivided into one of the three paradigmatic classes: root – the obligatory configuration – a single one per *F0* contour; prefix – the optional configuration preceding the root – there could be one or more of them in a contour; and suffix – the optional configuration following the root – non-recursive, i.e. only one in an *F0* contour. This classification reminds, though remotely, of the intonation phrase structure in the British school descriptions, with the nucleus viz. the root as the central element in the structure, and the prenucleus (prehead and head), as well as the tail, viz. the prefix and suffix, are thus the peripheral elements. The IPO structure of an *F0* contour also reminds to some extent of the autosegmental-metrical descriptions, where the corresponding pitch accents would be connected by the interpolating straight lines.

The quintessence of the phonological classification in the IPO model, based upon analysis of the common *F0* contours, were the prototypical pitch patterns. ‘t Hart and colleagues singled out six such patterns for the Dutch intonation, the most common ones being the hat, the valley and the cap pattern.

The IPO model inspired its further application to other languages beyond the original Dutch model. For British English, De Pijper (1983) implemented the model, following a similar basic procedure for identifying the perceptually relevant patterns. Thus, as a result of corpus processing without any prior phonological input, the scholar managed to extract eight basic intonation patterns and five tones. Tone 1, in all its four realizational variants, resembled the basic falling nuclear configuration; tones 2 and 3 represented different varieties of the rising pitch change; tone 4 would correspond to a fall-rise and tone 5 to a rise-fall in the British school description. De Pijper claimed that the rule-generated contours and tones were perceptually indistinguishable from normal speech. Similar conclusions were shared by Willems and colleagues (1988), who analyzed the stylized contours of British English pertaining to three pitch levels, i.e. high, mid and low, and with a varying rate of pitch change – abrupt or gradual. The resulting pitch sequences were "*highly valued by native listeners*".

Adriaens (1991) also used the IPO approach to construct his German intonation model, based on a corpus of read speech. Although the *F0* contours produced by the model, were rated as perceptually acceptable by the listeners, the corpus was very small to judge about the whole inventory of *F0* patterns available in German.

To summarize, the IPO model was a bottom-up approach to the description of intonation, gradually reducing the phonetic variation with the aim to single out phonologically, i.e. perceptually and articulatorily, relevant set of categories (the procedure is illustrated in Figure 4.4), corresponding to the basic *F0* movements.

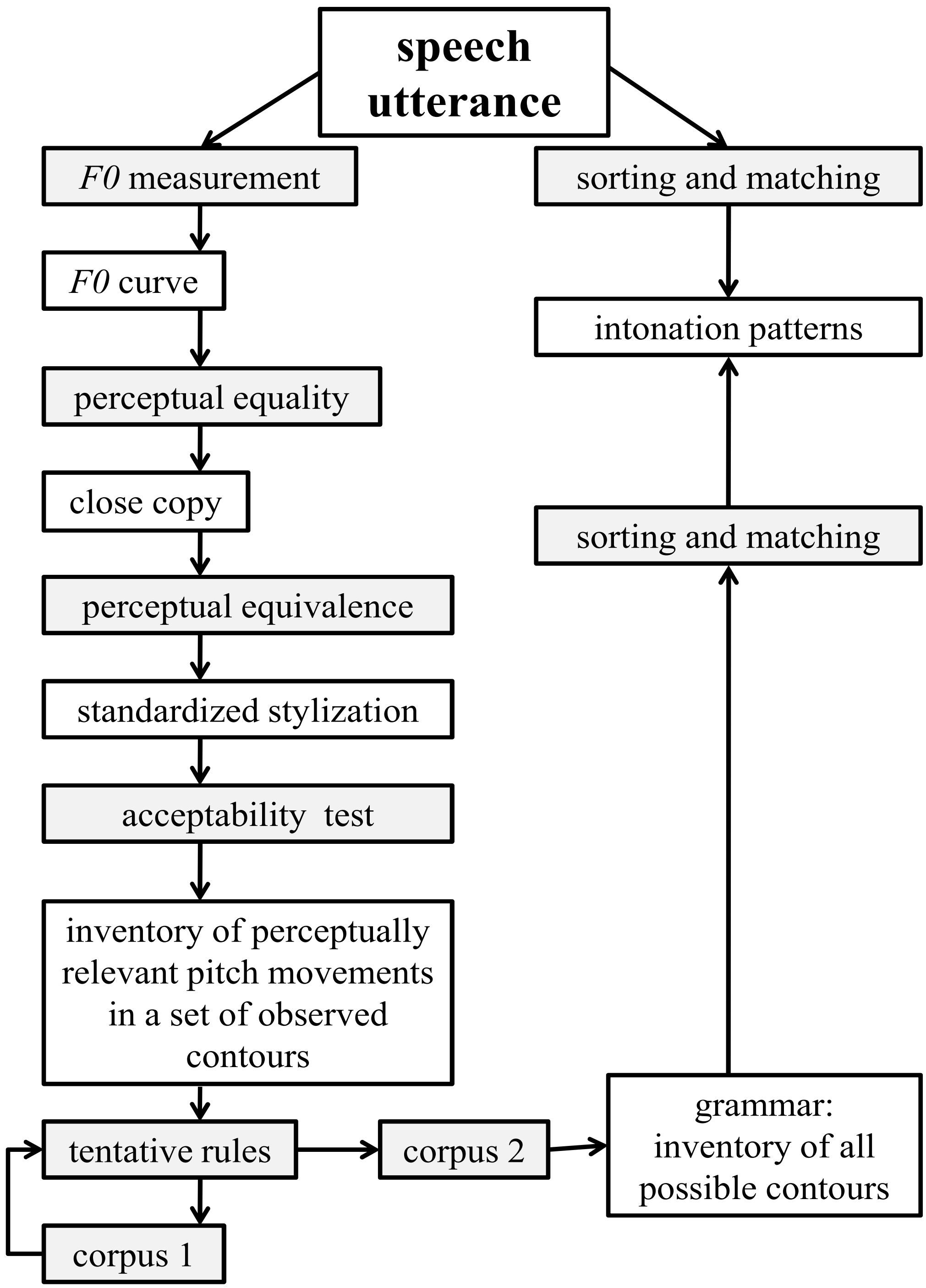


Figure 4.4: Summary of the IPO model process of prototypical contour extraction. The grey boxes represent processes; the white boxes – the products at each particular stage

4.3 Phonetic models

The theoretical and experimental-phonological models described in the previous sections provided comprehensive descriptions of prosody. However, no single paradigm is capable of accounting for the whole richness of intonational variation in speech, which may become especially problematic when those models are implemented in the modern technology – for example in automatic speech synthesis and recognition. With the aim to overcome this, a "third solution" emerged, and, namely, the idea to abandon the upper phonological level of analysis, operating on dynamic phonetic input.

Thus, phonetic proper models did not strive for any phonological classifications, though some of them do contain systemic relevance, even with no categories postulated. One group of such models is commonly referred to as superpositional.

4.3.1 Superpositional models

Superpositional models assume two levels of *F0* generation: the local events, approximately corresponding to pitch accents, and the global level, or the sentence prosody, whereas the latter would be superimposed on the former to create a functional *F0* contour. Although no phonological categories are postulated in these studies, the system is carved by the above functional building blocks per se, resulting in a basic hierarchical structure of a given intonation contour (Möbius 1993).

4.3.1.1 Öhman's original superpositional model

Öhman (1967) proposed the first well-known superpositional articulatory model, applied for Scandinavian intonation. The paradigm was built around the word and sentence intonation, which was further processed in word and sentence intonation filters.

The resulting *F0* contour was analyzed in terms of the basic macrounits – basic phrase contours, consisting of falling-rising movements.

Thus, the total *F0* configuration of the basic phrase contour was computed in a step-by-step manner:

1) assuming that all syllables are unstressed, a basic *F0* contour had to be calculated, as well as the contour boundaries;
2) the sequence of prominent syllables needed to be specified;
3) the energies of the prominent syllables had to be calculated. It should be noted that each successive stress was assigned a reduced stress value, as compared to the previous one;

4) the amplitudes of the phonatory stress pulses had to be calculated using the result of the previous stress. Apart from that, a stressed syllable would be lengthened in proportion to the energy of the phonatory energy placed on it. Öhman also speculated that it would be possible to introduce a number of primary stresses at the initial stage of contour generation and, further on, to assign the secondary accents based on the preceding primary stress. Or, alternatively, to introduce a prominence hierarchy between the primary stresses as well.
5) the basic phrase contour had to be recalculated and resegmented into syllables;
6) phonatory stress and word intonation pulses had to be introduced.

Next, the model would compute the pitch accent, i.e. sentence intonation, values as positive phonatory pulses, whereas the word targets would be represented by negative pulses. As a whole, the basic phrase contour was related to the breathing cycle, so a certain amount of energy was applied to the whole *F0* contour.

Although the author mentioned that the actual properties of *F0* contours had not received a sufficient analysis in the given study, it was still possible to single out three typical patterns of the basic phrase contour: the terminative mode, characterized by a steeply falling end contour, common for declarative sentences; the continuative mode, realized in a level-end contour, used for non-finality expression; elicitative mode, with a rising-end contour, used in questions and to express surprise.

The functioning of the Öhman model is illustrated in Figure 4.5, and the developed superpositional models – in the subsequent sections.

4.3.1.2 Fujisaki's model

Fujisaki (1983) adopted Öhman's original model and constructed an influential formal framework for describing intonation.

First of all, Fujisaki introduced the phenomenon of declination into the model, which was one of the features of the *F0* superposition, i.e. the general pitch configuration would supersede the local events to some degree.

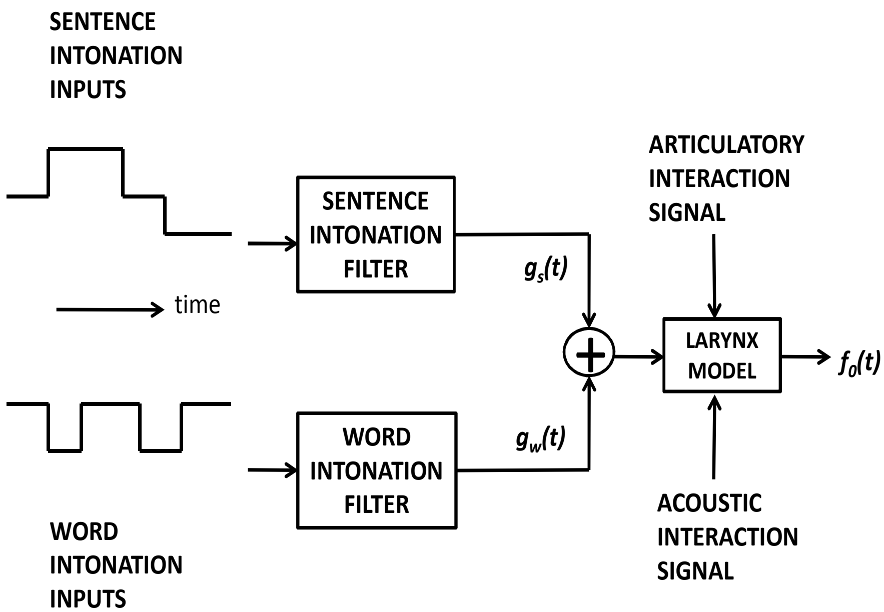


Figure 4.5: Öhman's model of speech production

Thus, a given *F0* contour represented a response of the phonatory system to a set of suprasegmental commands: the utterance command and the accent command. The former would produce the baseline component, whereas the latter would stand for the accent component of an *F0* contour. Both commands were considered to be discrete and binary, realized on a logarithmic scale.

Several variables were taken into account to compute the *F0* contour in the model: *F0* minimum for a speaker, amplitudes of the utterance and accent command, the onset and the end of the utterance command, the onset and the end of the accent command.

In general, a sentence *F0* contour displayed local *F0* peaks with a superimposed smoothly decaying baseline. Each local *F0* peak would correspond to a word or a word compound, and the peaks may be of equal or different *F0* height. The baseline component could be interrupted and resumed at major syntactic boundaries. The corresponding phrase commands were indeed assumed to take the form of a step function, but their amplitudes may not necessarily be equal for all the phrase components. A phrase command could be computed for each of those components.

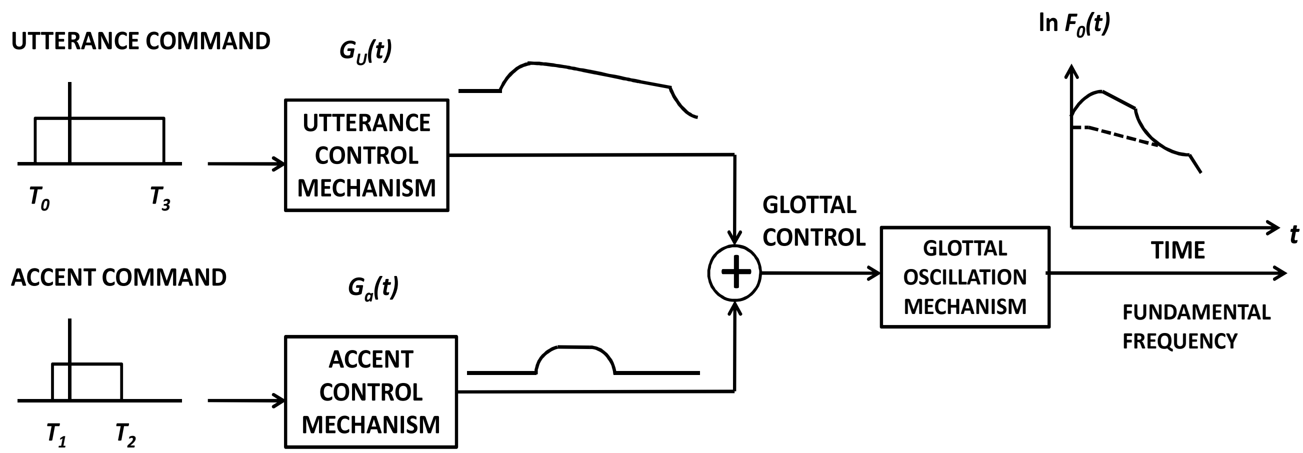


Figure 4.6: The illustration of the Fujisaki model

Möbius (1993) successfully applied the Fujisaki model to German intonation with a constant declination factor. However, similar models for English failed to produce satisfactory results for all intonational categories (Liberman & Pierrehumbert 1984) – the model failed to generate some patterns, e.g L*H and L* accents.

4.3.2 Unclassified phonetic models

Some phonetic models cannot be easily classified as pertaining to a certain paradigm. However, unlike the superpositional models, where some inherent phonological classification might be derived through their constituent parts, these approaches rely almost solely on low-level phonetic detail.

4.3.2.1 Van Santen and Möbius' quantitative model of *F0* generation and alignment

Van Santen and Möbius (2000) proposed a model, according to which pitch accent curves in the same class are generated from a common template using a common set of alignment parameters. These alignment parameters specified, how the time course of these curves depends on the durations of the segment sequence, with which a pitch accent is associated. The parameters in question were further embedded into a superpositional model, similar to that proposed by Fujisaki.

Thus, the main element in the model was the local pitch accent, representing a full pitch excursion for each event of this kind. Within a single curve, the authors paid special attention to the location of the *F0* peak, which was calculated based on the segmental structure of the respective syllable. In particular, the authors assumed, relying on some previous studies (e.g. Van Santen & Hirschberg 1994), that for monosyllabic groups peak location is systematically later in sonorant-final accent groups than in obstruent-final accent groups, and later in voiced obstruent-initial accent groups than in sonorant-initial accent groups. In polysyllabic accent groups, on the other hand, peaks occur significantly earlier.

Not only the pitch accent peak itself was of importance in the study: Van Santen and Möbius also singled out the so-called anchor points, i.e. the other important locations in an individual pitch accent.

As a result of analysis of different types of syllables, the scholars classified them into several types of templates, from which each individual pitch accent curve could be calculated.

For this purpose, a stepwise procedure was followed:

1) first, the pitch range was for a particular syllable type, additionally reflecting the prominence level of that syllable;
2) next, the pitch range would be scaled in order to calculate the average values for similar pitch accent curves, taking their location in an utterance into account
3) finally, special rules would apply for transitions between individual pitch accent curves.

At the final stage, pitch level of the accent curves, as well as the general contour of an intonation unit was computed. The former would be determined by a multitude of factors: position in prominence gradation, intrinsic pitch, whether a pitch accent were in a minor or a major phrase, etc. The phrasal curve would contain several anchor points, e.g. for English there would be three: the start of the phrase, the start of the last accent group in the phrase and the end of the phrase.

4.3.2.2 Taylor's RFC and Tilt Models

Taylor developed two models of intonation: RFC – Rise / Fall / Connection (Taylor 1995) and Tilt (Taylor 2000), which were approaches to the analysis of the surface acoustic patterns observed in *F0* contours.

In the earlier RFC Model, such patterns were said to pertain to two main types: peak (high) and trough (low) accents. According to Taylor, the former would roughly correspond to H*, H*+L¯ and L¯+H* accents in Pierrehumbert's (1980) model and to falls, rise-falls and high-level stressed syllables of the head in the British school formulation. The latter, in their turn, should be equivalent to L*, L*+H¯ and H¯+L* pitch accents, as well as low stress and the rising nuclear tone in the British descriptions.

The author, however, stressed the fact that even if it were possible to correlate peak and trough accents to the existing theoretical models, his approach was not intended to present any phonological classification.

The peak accents were modeled by a monomial equation as a combination of a rise element followed by a fall element. The gradients of these rises and falls were defined by a set of parameters. The trough accents, in an opposite manner, were derived from a similar function, but as falling-rising sequences of elements.

The boundary tones, both initial and final, were only modeled by a rise element, due to the frequent presence of sharp upward excursions on phrase edges. Low boundaries were not modeled separately. Following the logic of the approach, they would be inherent in the peak accents configuration.

Alternatively, they could also be classified as a connection element, described as an area in an *F0* contour, where nothing of interest would occur, i.e. no pitch dynamics could be detected. In such cases a straight line would be used to model *F0*, varying in duration and amplitude, i.e configuration as well.

Taylor undertook some experiments to test the paradigm in question. More specifically, a speech corpus was annotated and further resynthesized using the same labels. The results appeared to support the suitability of the model for speech synthesis purposes. As mentioned by the author, it could also be used for phonological descriptions as any higher-level categories can be derived from the basic model elements.

The Tilt Model was an improved modification of the RFC Model, which was achieved through the application of several underlying parameters: duration, amplitude and tilt itself. Duration was seen as the sum of the rise and fall durations. Amplitude was the sum of the magnitudes of the rise and fall amplitudes. The tilt parameter was a dimensionless number, which expresses the overall shape of the event, independent of its amplitude or duration. It was calculated by taking the ratio of the differences and sums of the rise and fall amplitudes and durations. According to Taylor (2000) the tilt representation is superior to the RFC representation in that it had fewer parameters without significant loss of accuracy. Taylor also argued that the tilt parameters are more linguistically meaningful.

4.3.2.3 Möhler's parametric intonation model: PaIntE

The PaIntE (Parametric Intonation Events) model is distinct from the other phonetic models in a way that it is both purely low-level and data-driven and, additionally, due to Möhler's preliminary considerations, it fits into the general autosegmental-metrical framework, especially the Tone Sequence Model (see Subsection 4.2.1 for a detailed account of autosegmental intonational models) and the German ToBI convention (Grice & Benzmueller 1994; Mayer 1995).

Möhler's (1998) original research, laid down in his dissertation, was based on a single-speaker corpus of German, where intonation contours were analyzed with the aim to model them using a higher-order mathematical function. Möhler considered linear interpolations between prominent points in an intonation contour to be rather crude approximations, so he developed an algorithm for a much more precise imitation of the natural *F0* excursions.

Though originally specified for German intonation, Möhler's model can be transferred to other Germanic languages, due to their relative similarity of *F0* excursions. In his own dissertation, the researcher also did some experimenting with the Lithuanian intonation using the same function.

In the original version of PaIntE, Möhler defined five basic parameters:

- ***l*** specified the *F0* minimum; ***h*** defined the pitch range of an individual curve; and ***l* + *h*** indicated the *F0* maximum;
- ***p*** modeled the shape of an *F0* curve: with the value of 1 it transformed the unit into a simple rise and with -1 – into a fall; rise-fall was constructed by applying 0 to the ***p*** parameter. Fall-rise was not included into the original model, due to the fact that it was not part of the German ToBI convention (for a detailed description see Subsection 5.2.1);
- the alignment of the function within a two-syllable window was defined by parameter ***d***. For a peak located in the exact middle of the stressed syllable, the parameter would take the value of 0, whereas negative or positive values would indicate a shift of the peak to the left or to the right, respectively.
- parameter ***s*** described the time, over which the function reached the *F0* maximum.[13]

[13] For the model formulae and illustration, it is advisable to consult Möhler's dissertation (1998) – pages 122-124.

The *F0* approximation itself was a step-by-step procedure. First of all, the original *F0* contour underwent a smoothing process, i.e. the voiceless and creaky parts were interpolated using the Secrest and Doddington (1983) algorithm. After that, the syllable boundaries were set using automatic alignment software (Rapp 1998). Each window contained two syllables, with a stepping interval of one syllable.

As the next step, Möhler's model defined the starting values for the function implementation. Thus, parameter ***p*** was determined by subtracting the values of the *F0* minimum from the *F0* maximum: should this be a positive figure, then ***p*** acquired the starting value of 0.5; in an opposite case ***p*** was ascribed -0.5; in cases with no significant difference between the two extrema, the parameter at issue equaled 0. The location of the *F0* maximum indicated the value of the ***d*** parameter, and the regression line to the extremum set the parameter ***s***.

The *F0* approximation employed the Nelder-Mead simplex search algorithm, in order to calculate the weighted error between the model and the original contour and to minimize it accordingly. If the error exceeded the acceptable limit, the model parameters were reset with new starting values.

PaIntE could quite robustly model the dynamic intonation events included in the German ToBI, i.e. falling, rising and rising-falling contours. Plateau, or valley, accents (e.g. H* or L*) would easily fall under the model definition, as they do not contain a significant pitch change – thus, no peak per se. Therefore, such units were modeled using a one-syllable window. In general, however, this approximation method was only employed when the approximation error went beyond 6 Hz. Otherwise, the two-syllable window was always taken as the default.

The evaluation of Möhler's original model demonstrated an 80% goodness fit of the data for the two-syllable window analysis. About a fifth of the data could be approximated with the single-window approach, and only 2% of the data could not be modeled with PaIntE whatsoever. Those would be syllables containing laryngealized, or the so-called creaky, regions, which commonly present problems for any kind of acoustic analysis. In general, the analysis of the original and approximated contours in the whole corpus under investigation resulted in a minor RMSE deviation of 6.8 Hz and a correlation coefficient of 0.94, which presented an improved outcome, as compared to some other studies (e.g. Black & Hunt 1996; Portele & Heuft 1997). Möhler also carried out a perception test for the contours produced by his *F0* approximation model, and these would more closely approach the original contour, than the curves resulting from neural-net training.

Möhler further developed a second version of PaIntE (Möhler 2001), changing the basic parameters, as well as the source model formula, and enriching the model with some new functionality. Thus, the ***p*** parameter was eliminated from the approximation process along with the ***l***. Instead, there appeared to be just one parameter indicating the *F0* peak in Hertz – ***d***. The other elements were modified to a varying degree: ***d*** in the original model was just renamed to ***b***;

s was split up into two separate parameters modeling the steepness of rises and falls separately, i.e. ***a1*** and ***a2***. These parameters were actually normalized. In order to derive their absolute values, they had to be multiplied by the corresponding values of two new parameters, ***c1*** and ***c2***, indicating the pitch range of the rising and the falling sigmoid. Like the ***d*** parameter, these latter were measured in Hertz.

The function below illustrates the underlying changes to the model, where the six new parameters find their place, as well as the constant alignment parameter γ.

$$f(x) = d - \frac{c1}{1 + \exp(a1(b - x) + \gamma)} - \frac{c2}{1 + \exp(-a2(x - b) + \gamma)}$$

The length of the analysis window was extended to three syllables, spanning over the prominent syllable, the one preceding it and the one following (see Figure 4.7), unless there were an interrupting boundary, in which case the window were cut off at the silence.

Figure 4.7 demonstrates a case, where all the six model parameters can be easily derived from the *F0* curve, i.e. the rising and the falling elements are present with sufficient prominence, as this would the case, for instance, with the rising-falling pitch accent (L*HL) in German. There is, however, a substantial number of other cases, in which one of the elements is either almost or completely non-existent: namely, the falling element in a rising tone (L*H), the rising sigmoid of a fall (H*L) or both the rise and the fall in a plateau accent (L* or H*). In such cases rises would be modeled with just the rising sigmoid of the function; falls, accordingly, - with the falling half-curve.

Mathematically, that was achieved by setting one of the ***a*** parameters to -1, whereby the corresponding ***c***-element would acquire the value of 0, if either ***a*** were less than 0.1 or ***c*** lay below 2.

Additionally, as described previously, the window for the simple-target accents would be reduced to the stressed syllable during the parametrization. A one-syllable window would also be employed for modeling boundary tones.

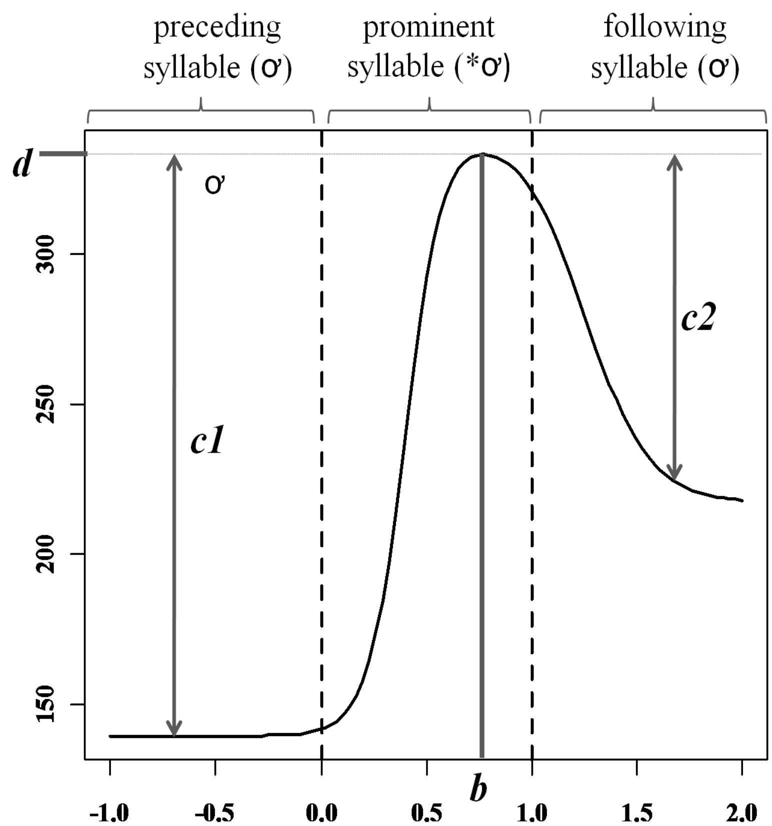


Figure 4.7: The PaIntE model function reproduced from Mohler and Conkie (1998). The F0 values are plotted on the y-axis (measured in Hertz), and the syllable-normalized time course – on the x-axis. The window covers three syllables; the syllable boundaries are graphically represented by dashed lines.

Still another modification to the model was the possibility to normalize the pitch range values. This option was added to account for the great variability of the phenomenon in question from utterance to utterance within one speaker and across individuals. The normalized ***d*** parameter would then vary within the range from 0 to 1. The disadvantage of this transformation is obviously the fact that global *F0* contour peculiarities, such as downstep and declination would be disregarded.

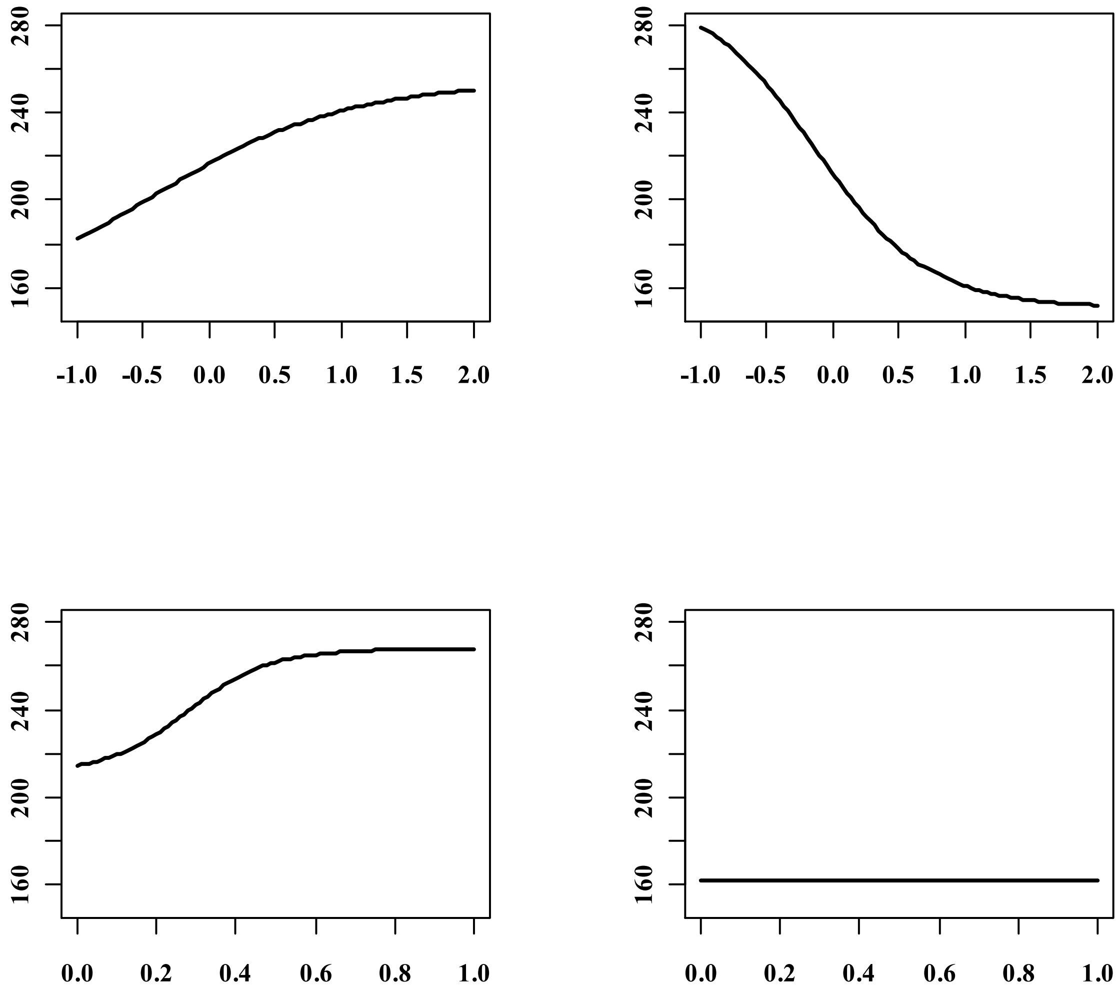


*Figure 4.8: Examples of single-sigmoid function implementation in various types of pitch accents: rising L*H (top left), falling (top right), simple high H* (bottom left), simple low L* (bottom right)*

Finally, Möhler applied a non-linear approach to *F0* peak placement. Thus, he relied on Van Santen and Möbius' (2000) idea of anchoring points in an *F0* curve (see Section 4.3.2.1 for a discussion) placed in different parts of the syllable. Consequently, each syllable was divided into the following constituents: the onset, mapped to the interval from 0 to 0.5 in the analysis window; the sonorant nucleus, occupying the interval from 0.5 till 0.8; and the coda, spanning from 0.8 to 1. The reason why the onset was granted the longest segment in the window was the fact that according to some investigations (e.g. House 1997), this element of the syllable plays the crucial role in the perception *F0* movements. The coda, on the other hand, is least influential in that respect. The anchor points of the pre-accented, as well as the post-accented, syllable were determined following the same logic as for the prominent syllables. Mapped were also the parameters ***a1*** and ***a2***, representing the length of the rising and the falling sigmoids, respectively, measured in seconds. These values were derived from the non-mapped parameters using the equation

$$a_i = 2_\gamma / a_i$$

The results of the experiments Möhler carried out to test this new anchoring division of the analysis window yielded significantly better results, i.e. lower RMSE and a higher correlation to the original contour, than the simple syllable normalization and syllable normalization coupled with pitch range normalization.

Some recent studies have employed the PaIntE model for investigating various phenomena. Particularly interesting in this respect is the attempt to predict ToBI categories from the raw model parameters, i.e. the actual opposite of Möhler's procedure. Schweitzer and Möbius (2009), who undertook this experiment, came up with considerably high precision rates for pitch accent and boundary tone identification in German and American English, comparable to and exceeding the results of other similar studies.[14]

Walsh et al. (2008) looked into the role of frequency of occurrence in the production of various pitch accents, attempting to study the similarity / dissimilarity between frequent and infrequent events. This was done by comparing individual *F0* curves represented by the six PaIntE parameters.

The model was also used to study pitch accents with regard to their information status by Schweitzer et al. (2009a, 2009b), who demonstrated significantly different realizations for new and given elements using PaIntE parametrization.

Summarizing the approach at issue, it is important to say that PaIntE provides rich phonetic detail in the form of the six basic parameters for each individual *F0* curve. Such an analysis may be carried out both with the underlying phonological classification, i.e. ToBI, and without it, looking at the raw *F0* contour, as was done in Schweitzer and Möbius' (2009) study mentioned above, as well as a more recent dissertation by Schweitzer (2011). Whatever method is chosen, it yields all-round information on a local *F0* contour: the alignment of its peak, with the possibility to anchor the location based on the syllable structure; the pitch range of the local excursion as a whole and as a sum of its rising and falling components; and, finally, the time course of the rising and falling sigmoids.

[14] For exact figures it is advisable to consult the original paper (Schweitzer & Möbius 2009) – page 2518.

On the whole, PaIntE is one in an array of *F0* parametrization approaches, which do not deal with real pitch configurations. Instead, a given intonation contour is modeled, or approximated. The advantage of the PaIntE Model is the fact that, despite approximation, an *F0* event is only stripped of minor phonetic detail imposed by the segmental context, the important pitch detail remains intact. Supporting this observation, Möhler's (1998) auditory tests, attest to greater perceptual acceptability of PaIntE contours, as opposed to the original natural *F0* contours.

4.4 Towards a model for cross-linguistic analysis of intonational variation

Each of the theories we reviewed in Chapter 4 can be applied to the description of German and English intonation. However, these models vary in their employability for exploring intonational variation from the cross-linguistic and SLA perspectives.

The IPO approach ('t Hart & Cohen 1973; 't Hart & Collier 1975; 't Hart et al. 1990) aimed at eliminating the phonetic detail and reducing variability to arrive at a set of prosodic prototypes. Whereas it would be an interesting task identifying prototypical contours in second language speech, it is equally important to investigate their concrete phonetic representations, in order to study the levels of deviation from the prototypes. This task would be beyond the focus of the IPO models.

The superpositional approaches (Öhman 1967; Fujisaki 1983) do inherently imply some level of prosodic variability when generating an *F0* contour and taking positional and other factors into account. These models, however, were mainly designed to generate intonation and not to analyze real-life corpora.

Van Santen and Möbius' (2000) model could potentially be used for analyzing some prosodic variation, e.g. pitch range variability of pitch accents and syllable types, as well as the temporal alignment of anchor points across speakers.

With Taylor's Rise/Fall/Connection Model (Taylor 1995) and Tilt Model (Taylor 2000) phonetic variability can be explored. However, phonological analysis is not foreseen in these models.

Möhler's (1998, 2001) parametric model PaIntE is capable of describing phonetic variation in terms of several parameters, which may or may not be linked to phonological categories – ToBI.

INTSINT (Hirst & Di Cristo 1999; Hirst et al. 2000), on the one hand, and the autosegmental models and ToBI, on the other hand, provide both phonetic and phonological description. The main advantage of INTSINT is that does not contain a predefined inventory of intonational categories, which gives the possibility for examining a given language and identifying a set of prosodic units for it. This, in turn, is an excellent device for studying underexplored languages.

If, on the contrary, one has to analyze well-explored languages, for example English and German, one might save the effort of creating a completely new prosodic system. It appears more rational to analyze the existing descriptions and modify them to suit a particular research purpose, if necessary. As described in Section 4.2.1.4, there are ToBI conventions available for a number of languages, including English and German, which provides a starting point for a cross-linguistic investigation of phonological variation.

The same holds true for the structural approaches to German and English intonation: they render ample phonetic and phonological information on the two prosodic systems. The main principles of Structural Phonology can also be transferred into other models, in order to supply additional focal points for analysis.

Summarizing the advantages of all of the investigated prosodic models, it appears that the most attractive approach for cross-language analysis of intonational variation is the one that allows to examine intonation from top to bottom: from the phonological level to the fine phonetic detail.

No single model that we have reviewed so far is capable of such an all-round description of intonational variation. Nevertheless combining two compatible models, ToBI for the phonological level and PaIntE (Möhler 1998, 2001) for speech phonetics, we could arrive at a two-level account of intonational variation.

Chapter 5 will discuss the concrete implementation the PaIntE and ToBI models in the given investigation, as well as describe the procedure of experiment.

Chapter 5

Methodological premises and procedure of experiment

The experimental part of the current investigation consisted in collecting a corpus of bilingual data, English and German, and further exploring the results of this effort based on certain methodological premises, which were viewed as applicable to such data. We will further provide a detailed description of how the study proceeded in this respect.

5.1 Experimental procedure

The current experiment involved several important components: subjects who agreed to participate in the study, their speech realizations of a certain kind constituting the corpus and the methodological basis, both theoretical and instrumental.

5.1.1 Subjects

The subject pool consisted of 41 native German speakers, 17 males and 24 females. At the time of the experiment the majority was aged 20-29, and six speakers were between 30 and 40 years old. The test persons predominantly came from the Swabian region of Germany and spoke standard German with a slight regional dialect.

All subjects learned English at school and most of them, except for three people, visited an English-speaking country, either sporadically, during short trips abroad, or over a longer period of time. The latter was true about a third of the informants.

A comprehensive series of tests had been performed by the subjects in the course of the DFG funded "Language Talent and Brain Activity" project (see Jilka 2009 and Jilka et al. 2007 for a detailed account of the procedure), prior to the given experiment. As a result, the speakers were subdivided into groups based on their pronunciation aptitude. We reported some results on those subjects before the current investigation (Anufryk 2009; Anufryk & Dogil 2009; Anufryk 2008; Anufryk et al. 2008). It should be noted in this regard that we applied a slightly different aptitude classification, as opposed to these previous studies. There, we used three-group analysis, i.e. the test persons were split into three main groups based on Matthias Jilka's classification – the below-average (proficiency scores 3.0 and below), average (scoring 2.0-3.0) and the above-average groups (scores 1.0-2.0). However, Jilka's evaluation allowed of a more detailed subdivision of speakers, which we decided to employ in the current study, in order to see whether the broader groups are really homogeneous or if some of them do not quite fit within such a general aptitude group. Thus, we defined six aptitude groups, corresponding with a more detailed gradation of Jilka's overall classification:

- highest –aptitude (3 m, 5 f) – speakers defined as exceptionally talented, scoring 1.0;
- high-aptitude (3 m, 4 f) – promising talented candidates, scoring between 1.1-1.5;
- above-average (2 m, 1 f) – subjects defined as "above-average", scoring between 1.6-2.0;
- average (5m, 7 f) – speakers exhibiting average performance, scoring between 2.1-2.5;
- below-average (1 m, 3 f) – subjects marked by degrading productions as opposed to average performance, scoring 2.6-3.0;
- low-aptitude (3 m, 4 f) – test persons certainly demonstrating deficiencies in L2 speech production, scoring under 3.0.

As seen from the above figures, most groups had a comparable number of subjects, with the average test persons having a slightly bigger representation in the corpus. The above-average and the below-average groups, however, each comprised just one speaker of one of the two genders. In cross-gender analysis this would not be considered, strictly speaking, a group. Nevertheless due to the fact that this study is concerned with the exploration of the prosodic variation phenomenon, the preservation of such data in the analysis would, among other things, reveal not only group tendencies, it would show individual variability as well. Although we believe that any such results cannot necessarily be representative of a larger number of speakers with similar pronunciation aptitude, they may yield interesting insights into the phenomenon in question in general.

In the earlier studies, the first three groups were defined as above-average, i.e. the highest-aptitude, high-aptitude and above-average speakers; the average group remained unchanged; below-average and low-aptitude speakers were collapsed into the below-average group.

The majority of test persons produced L2 English speech somewhat in-between the British and American variety and with an influence of the German phonetics and prosody. Only several of the highest-aptitude speakers produced speech perceptually very close to the native realizations. More specifically, there was one female subject possessing a clearly British accent, and one female and one male speaker approaching the L1 American English pronunciation. These observations were based on judgments of three expert phoneticians. The rest of the subjects produced greater degrees of Germanness in their English samples.

For the validation of all the L2 tendencies, a control group of twelve native English subjects (4 females and 8 males) took part in the experiment as well. The majority were speakers of American English, with the exception of one Irish male national and a British female test person.

However, with regard to variation patterns, these speakers did not impact the experimental results, which was supported by statistical tests, and their productions were therefore preserved in the corpus.

The participation of all subjects was financially and respectfully appreciated.

5.1.2 Corpus and data processing

The subjects described in Subsection 5.1.1 were recorded in the anechoic chamber of the Institute for Natural Language Processing, University of Stuttgart, with a sampling rate of 48 Hz to allow of high quality in the realizations.

The experimental task consisted in producing both read and spontaneous speech. For the first speech type, the subjects read two abstracts: the classical fable “The North Wind and the Sun” and an abstract from Mark Twain’s “Mrs. McWilliams and the Lightning“. The first text was rather neutral in nature, whereas the latter one contained some dialogue and therefore the potential for some expressivity in realizations. As the second task, the test persons were supposed to speak spontaneously, or rather, quasi-spontaneously, in an interviewer-guided conversation following a similar set of questions for each subject. They had to mention the following points:

- their name and origin;
- their education and occupation;
- how long they learned English;
- their self-evaluation of language skills;
- if they know any other foreign languages;
- their trips / experiences abroad;
- their hobbies and future plans.

The questions were presented in the same order to each of the informants.

The German subjects performed their read and spontaneous speech tasks in both German and English. The native English speakers performed these tasks only in their mother tongue English.

The recordings were further downsampled to 16 Hz/16 Bit for technical processing reasons. Word, syllable and phone boundaries were marked automatically by forced alignment using the Aligner software for German and English (Rapp 1998) and corrected manually for any resulting errors.

Manual labeling of the intonation events followed, in accordance with the general principles of autosegmental metrical phonology and the ToBI convention, but with certain modifications, seen as indispensable in the light of the present investigation, which we will describe in Section 5.2.1.

The syllables carrying ToBI accents and boundary tones were further analyzed in a parametric intonation model (Möhler 2001), introduced in Chapter 4 and described from a detailed technical perspective in Section 5.2.2.

5.2 Methodological premises

5.2.1 Annotation approach

There exist a variety of approaches to the description of intonation in any given language, and there is a number of annotation frameworks to support them. We described models of German and English intonation in Chapter 4. Among others, we spoke in quite some detail about autosegmental-metrical phonology and the ToBI convention emerging from it. We then also suggested that, other paradigms available, ToBI appears to be especially suitable to the current study due to the fact that there are descriptions of both German, in its Stuttgart and Saarbrücken versions, and English, British and American, in this respect. Whereas they do postulate certain inventories of pitch accents and boundary tones, it would be interesting to see how these categories are represented in real speech.

The challenge for cross-language comparisons is, however, that the English and German annotation conventions exist separately, and with some discrepant underlying theoretical premises resulting in differing symbols for similar events. Secondly, L2 speakers' realizations vary in the degree of nativeness / accentedness of target language categories and can be therefore difficult to classify using an L2 inventory. Apart from that, there may be a considerable amount of L1 transfer, which means that German subjects would map the typical intonation contours from their mother tongue onto English.

For all of the above reasons, we decided to adopt a unified approach to the annotation of intonation events in the current study. Namely, the best-suited pitch accent and / or boundary tone from all the available ToBI inventories was selected for each individual token.

This was done in order to facilitate the categorisation of non-native speech samples as complying or failing to comply with a certain L2 target. The resulting distribution and frequency of intonation events would then reflect an L2 speaker's performance in L2, and this, eventually, would serve as an indicator of his or her language aptitude. Thus, should a subject employ appropriate categories with a frequency of occurrence adequate to that of a native speaker, then this should point to a better proficiency and higher aptitude levels, and vice versa. Applying the same approach to native speakers' realizations, on the other hand, might help identify the basic and marginal variants within categories, as well as reveal new categories, as some tokens might be potentially difficult to classify within the existing system.

In the case with differing events in the two languages, all of them would be added to the general inventory. However, a number of modifications were necessary, in order to address the above-mentioned differences in the annotation of similar intonation phenomena.

5.2.1.1 American, British and German ToBI merged

ToBI was first developed for American English by Silverman and coworkers (1992); then one more version followed as a modification of the original system (Pitrelli et al. 1994). The British "ToBI" is in fact the "IViE Labeling Guide" (Grabe 2001), as it was created for annotating the respective corpus of British English regional dialects. Finally, there exist two versions of the German ToBI: one designed in Saarbrücken (Grice & Benzmüller 1994; Baumann et al. 2000) and the other in Stuttgart (Mayer 1995). As described in Section 4.2.1.4, the two German paradigms differ in theoretical premises: the Saarbrücken system followed the American English conventions rather closely; the Stuttgart version, on the other hand, integrated the autosegmental-metrical intonational analysis of German intonation done by Féry (1988, 1993), which departed from the original American English ToBI quite significantly. In the current study we will employ the Stuttgart ToBI, based on the preliminary expert analysis of our experimental corpus: the individual tokens could be classified more easily within this version of ToBI, thus suppoting the core theoretical assumptions behind it (see Section 4.2.1.4 for a discussion).

There exists a considerable degree of agreement between the German, British and American English ToBI. Essentially, they describe the same types of intonation events: a) "static", i.e. steady-pitch events, or plateaus, represented by monotonal pitch accent and boundary tones – simple H (high) and L (low) targets applied to accented syllables and phrase edges; b) "dynamic", i.e. involving a significant pitch-change – rises and falls, or combinations of rising and falling movements – a sequence of H and L events. Of course, that this division into dynamic and static events is but relative, as there are slight changes in pitch even within the static pitch accents and boundary tones. What is meant here is the presence or absence of a significant pitch alteration. We did introduce the two types of intonation phenomena because they are at the core of the annotation approach in the current investigation meant to resolve the conceptual differences in the description of English and German intonation within the framework in question.

Boundary tones

In American ToBI (Pitrelli et al. 1994) each utterance is described as consisting of full intonation phrases (IP's) and intermediate phrases (ip's), i.e. sense blocks split by longer and shorter pauses, respectively. At minimum, an utterance should consist of one ip, which, in turn, would constitute an IP, as ToBI presupposes a hierarchical structure of intonation units, where an intonation phrase spreads from the left edge of the first intermediate phrase till the right edge of the last ip. Thus, in case of single-ip intonation phrases, intermediate phrase boundaries coincide with IP edges. Normally, however, an utterance contains several IP's, each possibly, but not necessarily, comprised of several ip's.

In notation, intermediate phrases are marked by the minus sign (-), whereby the percentage sign (%) signifies an intonation phrase boundary. It should be noted in this respect that the IViE version only includes full intonation phrases in its description. In the current investigation we consider intermediate phrases to be an important part of utterance structure and therefore preserve the American English classification, as some discourse chunks can hardly be annotated as fully-fledged intonation phrases. Instead, there are often slight hesitations and / or pauses in speech, which mark minor boundaries, then followed by a more prominent break, complying with the notions of the intermediate and intonation phrases, respectively.

According to the pitch level, boundary tones may only be of high (H) or low (L) variety in the original American ToBI system. However, the German ToBI and IViE introduce a third type, namely, the "default" boundary, labeled as a single percentage sign (%) for full boundaries and as a single minus sign (-) for ip's – the latter only exists in the German notation. This edge type stands for IP's / ip's with no significant pitch change after the last pitch accent.

It is true that any intermediate or intonation phrase will either end on a low or on a high *F0* level. However, that would not establish any relationship with the previous pitch accent: did the contour continue to rise after it or drop or remain unchanged? It apprears that the notion of a default boundary helps define such a relationship. Thus, if we reserve the H and L boundary targets for the above-mentioned "dynamic" events, i.e. for a falling or rising pitch configuration until the boundary, then the default edge type would mark cases of almost straight-line, unchanging *F0* segments before an IP / ip break. One might argue that American English ToBI does contain such a boundary tone already, which is H-L% and its downstepped variant. We shall, in fact, use it in appropriate cases in the corpus, along with the other categories. However, this boundary type presupposes a slight *F0* lowering, marked by the H / L interpolation. In the pure default boundary, on the other hand, no pitch change is to be detected. Our prior empirical examination of the corpus confirmed the fact that such intonation events, default boundaries, are rather common, especially in spontaneous speech. One might also argue that default boundary events differ in pitch depending on the previous pitch accent: there will be a high pitch level after a pitch accent ending in a H tone and low pitch will follow an L tone. In the given experimental corpus, the majority of default boundaries were to be found after H tones; only a marginal number of such tokens was found after L targets. This fact can be explained by the mere physical difficulty to pronounce stretches of speech on a low unchanging pitch. Therefore, those rare instances were excluded from the statistical analyses.

In general, the inventory of boundary tones in the current investigation is a sum of all categories in British, American and German ToBI, which is presented in Table 5.1.

Boundary type	**Boundary tone type**	**Symbol**	**Description**
intermediate	dynamic	H-	*F0* rise after the last pitch accent in an ip
		L-	*F0* fall after the last pitch accent in an ip
	static	-	no pitch-change after the last pitch accent in an ip
full	dynamic	L-L%	*F0* fall to a point low in the speaker's range after a low ip boundary, e.g. in declarative utterances
		L-H%	*F0* rise after a low ip boundary, e.g. in continuation rises
		H-H%	*F0* rise to a very high value in the speaker's range after a high ip boundary, e.g. in general questions
		H-L% (!H-L% for the downstepped variant)	a slight *F0* fall after an upstepped high ip boundary to a mid level in the speaker's range, e.g. in calling contours
	static	%	no pitch-change after the last pitch accent in an IP

Table 5.1: Inventory of boundary tones and their description

Pitch accents

Pitch accents are associated with the most prominent syllables in an intonation or an intermediate phrase. Like the boundary tones, they are pitch level targets – high (H) or low (L): single ones represent static pitch accents with no significant pitch change on the accented syllable or in its immediate vicinity; apart from that, they might also be sequences of two or three simple targets associated with the prominent syllable, in which case dynamic events (falls, rises, fall-rises and rise-falls) are in place.

In the original American ToBI notation, the accented syllable in a pitch accent is marked by the asterisk sign (*); the plus sign (+) signifies the trailing tone before or after the respective accented syllable. For convenience and following the example of other studies (e.g. Mayer 1995, Ladd 1998), we will omit the plus sign notation, thus labeling the trailing tones simply as "H" and "L".

There is some agreement as to the actual inventory of pitch accents in the three ToBI versions. More specifically, two simple pitch targets H* and L*, as well as the "scooped" rising accent L*H, are equally common for American, British and German ToBI. The other pitch accents are language specific.

Thus, the rising LH* and the step-down H!H* events are only to be found in American English ToBI. However, the latter is similar to the German "early peak" accent HH*L, especially if a low boundary follows H!H*, forming the nuclear fall.

The IViE paradigm singles out the falling-rising H*LH accent, which is absent from the other two ToBI versions. In the American ToBI the ip-/IP-final fall-rise is usually marked by a combination of the H* accent and L-H% boundary. However, such a notation does not account for fall-rises in the non-final position in an ip/IP, as well as on monosyllabic words. The IViE label would then be preferable for those two cases and where the rising-falling coherently marks the prominent syllable and its immediate context and is not spread across a number of words. Worthy of mentioning is also the fact that the fall-rise has been mentioned as an important and frequent category in English intonation by a number of studies, especially the ones pertaining to the classical British school (e.g. Crystal 1969).

Both German and British ToBI include the falling H*L and the rising-falling L*HL pitch accent. American ToBI employs the combination of H* followed by a low boundary for the former and LH* immediately preceding a low boundary for the latter. In the current study, in these cases we will look first and foremost at the prominent syllable and the pitch change associated with it. Consequently, should the falling movement start within the target syllable, then it is to be labeled with an H*L accent, otherwise if there is no significant *F0* movement there, then the token is to be treated as carrying a simple high-level tone. For rise-falls the annotation procedure will be similar. If the accented syllable can be classified as an LH* event, i.e. a sharp pitch rise from a low level onto a high level of the starred syllable, then a low boundary would be a logical

Pitch accent type	Symbol	Description
static	H* (and the !H* for the downstepped variant)	a high target on the accented syllable not immediately followed by a fall
	L*	a low target on the accented syllable not immediately followed by a rise
dynamic	L*H (and L*!H for the downstepped variant)	"scooped" accent; a low tone target on the accented syllable which is immediately followed by relatively sharp rise to a peak in the upper part of the speaker's pitch range or a rising pitch-movement starting, or completely realized, already in the accented syllable
	LH*	"rising peak accent" – a high peak target on the accented syllable which is immediately preceded by relatively sharp rise from a valley in the lowest part of the speaker's pitch range
	H*L (and !H*L for the downstepped variant)	a falling pitch-movement starting, or completely realized, on the accented syllable
	H!H*	a clear step down onto the accented syllable from a high pitch which itself cannot be accounted for by a H phrasal tone ending the preceding phrase or by a preceding H pitch accent in the same phrase; should only be used when the preceding material is clearly high-pitched and unaccented (otherwise, the accent is a simple !H*)
	HH*L	"early peak"; a high target on the preaccentual syllable followed by a fall on the accented syllable. This accent type must be realized on at least two syllables: an accented syllable and a preaccentual syllable, which must be unstressed. The fall starts, or is completely realized, on the accented syllable.
	H*LH	fall-rise: high target (or a fall) on the accented syllable followed by a rising pitch-movement

Table 5.2: Inventory of pitch accents and their description

continuation of the rising-falling pattern. If, on the other hand, the rise-fall is compactly placed on 3-4 syllables, with a gradual rise on the prominent one and the low trailing part, then the target is to be granted an L*HL status.

By analogy with the approach to the final falls, the nuclear rises will marked as a sequence of L* and a high boundary, as in American ToBI, if the rise is not associated with the accented syllables and is somewhat independent of it, which should be manifested by the fact that there is no immediate ascending *F0* movement, following the pitch accent. In the contrary cases, i.e. with a rise closely following the starred target, the L*H is to be employed, which can be used in any position in an IP / ip.

The only pitch accent specific for German is the stylized "calling" H*M contour, which is defined as a slight fall to the middle of the speaker's pitch-range after an unchanging high pitch. This description concurs with that of the H-L% plateau boundary in American ToBI. The only phonetic difference seems to be the lengthening of the final syllable(s) mentioned in the German convention, which is also taken as a discriminator between the two notation variants in the current study.

Finally, it should be mentioned that both the basic pitch accents and boundary tones and their downstepped variants are used in the labeling process. The complete inventory of pitch accents and their description is given in Table 5.2.

5.2.2 *F0* parametrization

After the corpus was manually annotated by two expert labelers using the unified approach described in the previous section, each of the resulting prominent syllables and *F0* curves was processed in the PaIntE model, which we described at length in Chapter 4.

Hereby, we would only like to clarify some technical details associated with the application of the model. Thus, PaIntE allows of different settings for several of its parameters.

More specifically, it is possible to extract both normalized and non-normalized values for the *d* parameter, i.e. the full *F0* excursion in an individual curve. We applied the latter variant, because for the current investigation of intonational variation it was important to see the actual pitch levels and pitch range variability within the groups, along with the "pure" variation, for which we then calculated z-scores for each *F0* value per speaker.

The alignment parameter *b* also allows of several ways of extraction. First of all, as with *d*-values, there may or may not be normalization involved. Here, we decided for the normalized option because it then reflects the relative position of the *F0* peak within the analysis window.

Otherwise, single tokens just carry time point stamps in the acoustic signal, which is not very informative for any further statistical comparisons.

The other parametrization settings remained standard.

Having acquired both the phonological and phonetic properties of the corpus, we proceeded with data analysis on these two levels. Chapter 6 will therefore dwell upon phonological variation patterns, whereas the final Chapter 7 will introduce results of the phonetic investigation of PaIntE parameters.

Chapter 6

Investigation of phonological variation patterns

At the initial stage of experimental data analysis, we investigated the variation patterns of phonological categories, i.e. ToBI pitch accents and boundary tones, with two major objectives.

The first one consisted in obtaining a general idea of the distributional tendencies in our corpus. That was supposed to allow us to single out frequent and infrequent intonation events in view of their possible correlation with the following factors:

1) first and foremost, level of pronunciation aptitude, reflected in the talent gradation described in Chapter 5, namely, the subdivision of the subjects into highest-aptitude, high-aptitude, above-average, average, below-average and low-aptitude, based on their acquired level of proficiency;
2) speech type – read or spontaneous – and its relation to the phonological variability patterns in the aptitude groups;
3) gender – for the potential differences between male and female subjects of varying aptitude as they realized utterances of the two speech types.

The second objective only pertained to part of the data, i.e. read speech realizations. Here, we aimed at revealing the commonalities and discrepancies as to the general text interpretation in the speaker groups by comparing the prosodic transcriptions of the two text abstracts – "The North Wind and the Sun" fable and an abstract from "Mrs. McWilliams and the Lightning" (by Mark Twain). Some preliminary auditory work implied the idea that some utterances are produced by different speakers in much the same way intonationally, and we therefore wished to test whether this fact has anything to do with a subject's pronunciation aptitude. More specifically, the degree of variation as well as uniformity on this meta-level within the aptitude groups was of essence at the given stage of analysis.

The results of the two phonological examinations will follow in the subsequent sections, whereby we will first discuss the frequency of occurrence and distribution of ToBI pitch accents and boundary tones and, after that, present the findings in the global text interpretation in the speaker groups.

6.1 Tendencies in the distribution of ToBI pitch accents and boundary tones

The first stage of phonological analysis consisted in detecting the variation and distribution patterns of ToBI categories – pitch accents and boundary tones.

The German part of the corpus was investigated first, due to the fact that the subjects in the current investigation are native speakers of this language. Therefore, assessment of the German data was supposed to lay a foundation for further comparison with the L2 English data, namely, finding any instances of the second language processes we discussed in Chapter 2, e.g. transfer from L1 into L2, avoidance of some foreign language phenomena, reduced variability of L2 speech as compared to native speech etc. The study proceeded with the research of the English part of the data per se, in order to examine the issues in L2 speech just mentioned, as well as any other potentially emerging phenomena.

Thus, each of the following two sections (Section 6.1.1 and Section 6.1.2) contains descriptions of the distribution peculiarities of pitch accents and boundary tones in general and with regard to the three factors put forth in the introduction to the current chapter – pronunciation aptitude, speech type and gender. The procedure of analysis of ToBI categories consisted in analyzing the whole corpus in general first, followed by gender-specific and speech type dependent examinations. Finally, both the speech type and gender factors were applied to the data.

6.1.1 German language data

6.1.1.1 Pitch accents

General corpus analysis

The initial analysis of the distribution of ToBI pitch accents in the six German aptitude groups revealed some frequency of occurrence regularities, which did not contradict with several general patterns for German. In particular, the two most frequent categories registered for the current corpus were L*H and H*, concurring with the observations of some earlier studies (e.g. Mayer 1995). The corresponding percentages of these two pitch accents provide a more detailed insight into this tendency.

Thus, the percentage of L*H accents was highest for the average and above-average groups (over 35% of all accented syllables), followed by the high-aptitude, below-average and low-aptitude speakers (~33 %). The highest-aptitude subjects were exceptional in this respect, with only ~27% of all prominent syllables being realized with the rising pitch accent.

Almost equally frequent was the H* accent, claimed to be the default for German in the above-mentioned study by Mayer (1995). It was found in about 30-35% of all the prominent syllables in the German part of the corpus in all the groups. The above-average test persons, however, applied the category in question to the smallest amount of syllables – just around 27%.

The simple low target L* was common for about 15 % of the syllables under investigation, with the highest frequency of occurrence in the above-average group (~17%) and the lowest percentages in the low-aptitude and high-aptitude groups (~12%).

To the frequent categories in the corpus also belonged the rise-fall L*HL and the falling accent H*L, each accounting for about a tenth of the test data. In the former category, the frequency of occurrence was almost equal in all subject groups.

The rest of the ToBI pitch accents could be classified as infrequent, comprising less than 5 % of the data. Those were proper German accents, i.e. the downstepped variants of the simple high target and the falling tone - !H* and !H*L, respectively, as well as the stylized tone H*M and the HH*L. Apart from that, some of the events could be best labelled within the English language inventory (American and / or British variety) – e.g. as LH* and H*LH, which points out to some observable phonetic differences between the tokens, which might have been applied to the same established German category in a monolingual study (Figure 6.1 demonstrates the general-corpus findings in the distribution of pitch accents in all aptitude groups in the German language samples described above). Nevertheless such instances were not numerous and did not present, along with the other infrequent categories, any distributional trends. We will therefore not present any analysis of the infrequent categories either for the entire corpus or for the further specification of data by speech type and gender.

Gender-specific analysis

After analyzing the distributional peculiarities in the whole experimental corpus, we proceeded with investigating the possible effects of gender and speech type on the employment of pitch accents in the corpus. Analyzing the trends in gender / talent groups, we found some similarities as well as mismatches in the distribution of ToBI categories. First of all, this concerned the L*H accents.

Rising tones L*H

With regard to the consistencies, the rising accents were realized by male and female subjects in high-aptitude and low-aptitude groups with almost the same frequency of occurrence – covering ~31-33 % of all accented syllables. Still another example of steady performance was rendered by the average groups of both genders: average male and female speakers retained the highest overall percentages of L*H accents (~35-37%).

The lowest general percentage of the given accents common for all highest-aptitude speakers, without the account of gender, characterizing the whole corpus, appeared to be influenced mostly by the female speakers, for whom this tendency was preserved, with under ~25% against the average of ~33% in all the other female groups.

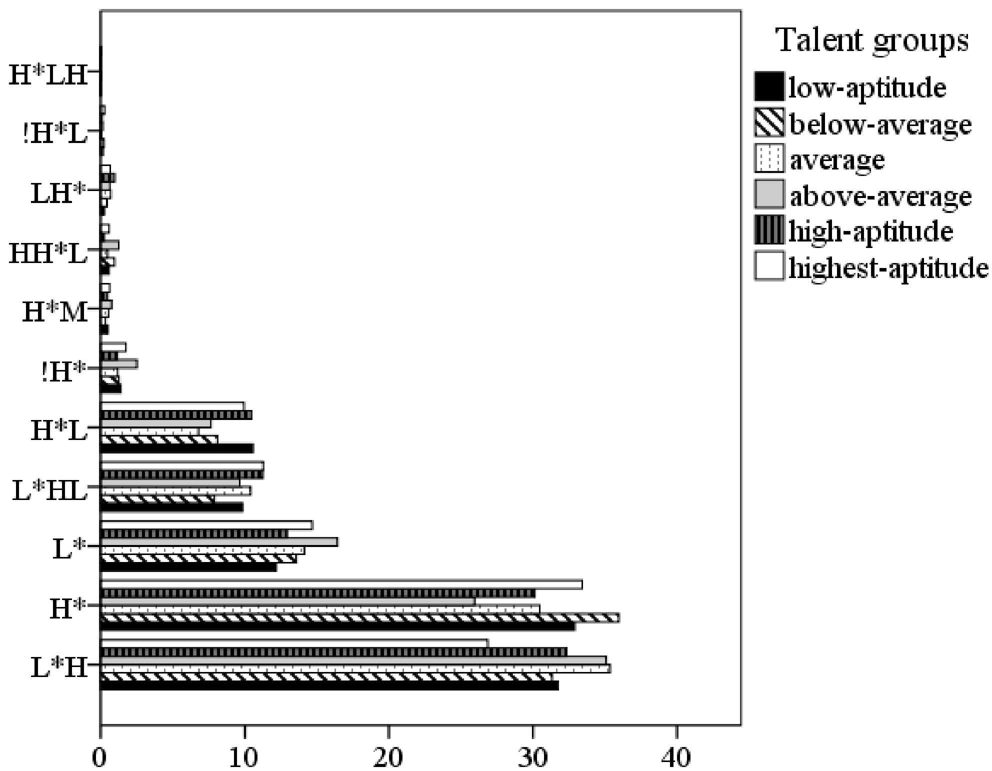


Figure 6.1: Distribution of pitch accents in all aptitude groups in the German-language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

The male subjects, on the other hand, replicated the pattern proper to high-aptitude and low-aptitude speakers in that respect (~32%). Finally, the above-average test persons exhibited some gender-specific distribution peculiarities as well: the female speaker followed the pattern set by the high-aptitude group, with ~33% of all prominent syllables being produced with an L*H accent, whereas male subjects clustered with the average subjects in that respect (~37 %).

The only below-average male subject produced his utterances in a highly idiosyncratic way, i.e. clearly different from the corresponding patterns of other male speakers, as well as the below-average female test persons. These differences included a significantly lower percentage of L*H (under 20%) accents than all the other male and female subjects. Without a doubt, single-subject distribution patterns cannot be interpreted as a general trend for male speakers with similar pronunciation abilities. The figures, which we have just presented, have to be tested on a larger scale. However, they do point out some individual variability patterns that contribute to group tendencies. In this sense they are therefore of certain value, as both the group and the individual factors are captured to some degree.

Simple high targets H*

The H* accents could also be characterized by both general and divergent tendencies between the aptitude / gender groups. Clearly divergent was the performance of two male groups in this respect.

The below-average male speaker, whose idiosyncratic behavior was mentioned above for the rising accents, also demonstrated a significantly higher frequency of occurrence in simple high targets (~43%). The above-average subjects, on the contrary, had the lowest amount of H* tokens (~25%), as opposed to the other male groups producing comparable numbers of the given pitch accents – on around a third of all prominent syllables.

With the female subjects as a whole, the H* accents enjoyed a slightly higher frequency of occurrence than in the male groups (except for the single below-average male outlier): it fluctuated from ~30% to ~35% in all groups.

On the whole, it should be pointed out that male and female speakers demonstrated comparable variability patterns in the H* accents, although there were some gender peculiarities in the distribution. The clear exception was made by the smallest groups, i.e. below- and above-average subjects, whose performance did not comply with the general trends quite observably.

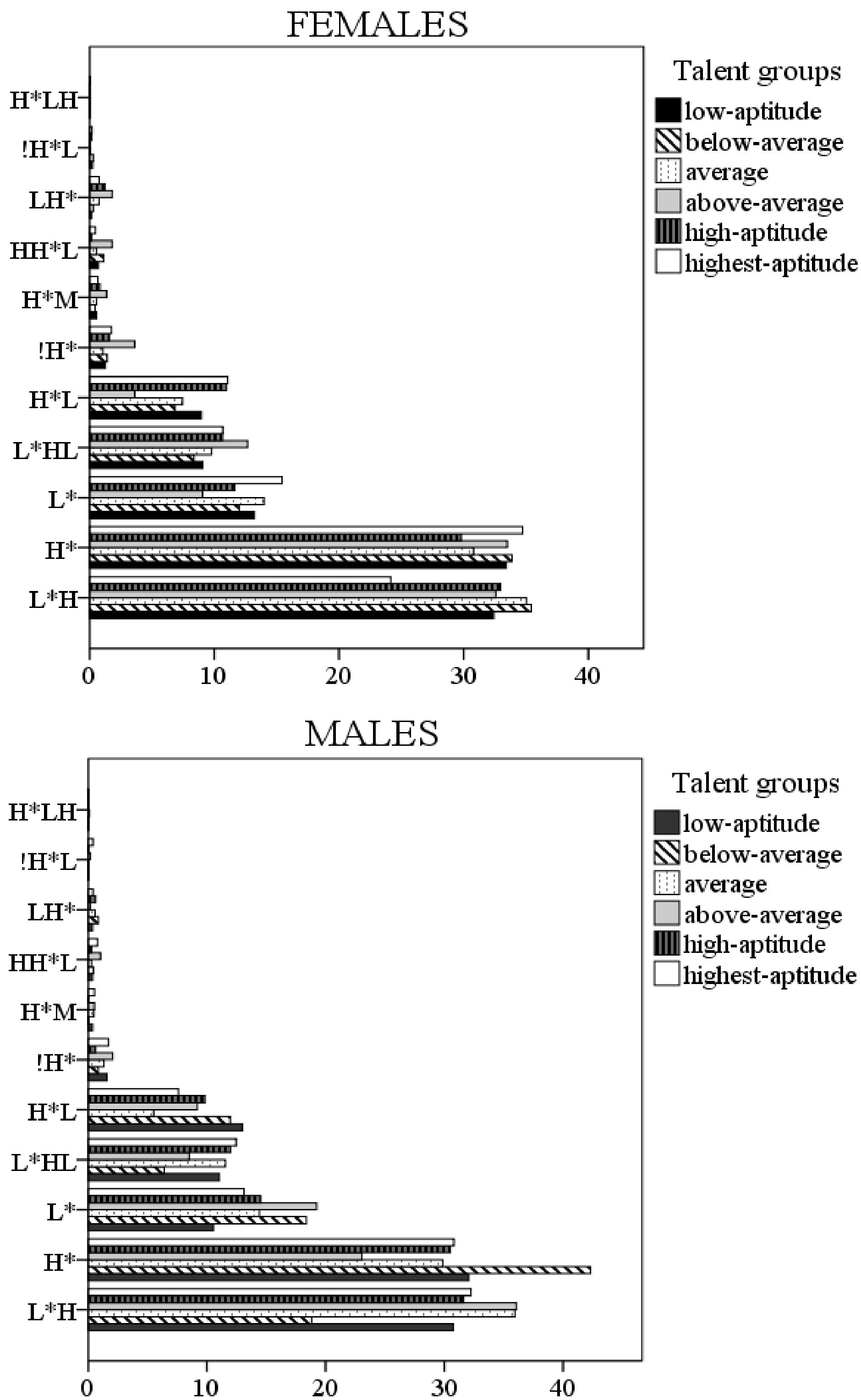


Figure 6.2: Distribution of pitch accents in all aptitude groups by gender – female versus male realizations in the German-language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

Simple low targets L*

As far as the simple low targets are concerned, their distribution was considerably regular as realized by both genders. The L* accents were applied to ~15% of all prominent syllables on average by male test persons and ~13-16% - by the female speakers. The only examples of deviating performance were presented by male below-average and above-average speakers (~20%), as well as the above-average and low-aptitude female informants (~9-10%).

Rising-falling tones L*HL

The average percentage of the rising-falling tones in male groups equaled ~13%. Below-average and above-average male subjects departed from this average values marginally, with under ~10% of L* targets in their speech.

Female speakers exhibited similar performance with respect to L*HL events by producing ~12% of such tokens on average. The above-average speaker employed the category in question slightly more frequently (~14%), whereas below-average and low-aptitude speakers, on the contrary, demonsytrated a generally lower frequency of occurrence of 8-10%.

Falling tones H*L

The final frequent category, the H*L accents, did not follow any of the regularities detected for the previously described pitch accents. Instead, the frequencies of occurrence patterns in this category were rather peculiar. For the male subjects, this category was most frequent in the low-aptitude and below-average groups (~13%); the female high-aptitude and highest-aptitude speakers demonstrated a similar pattern. The lowest number of the falling accents could be found in the average male (~6%) and above-average female (under ~5%) groups.

The remaining test persons of both genders could be located between the above frequency of occurrence extremes, applying the falling tones to up to a tenth of all accented syllables.

All gender-specific peculiarities of pitch-accent distribution in German-language productions are illustrated in Figure 6.2.

As stipulated earlier, the gender factor executed some influence on the distributional picture in all talent groups under investigation. Therefore, we also studied another potential impact factor in this respect, namely, that of the speech type.

Corpus analysis by speech type

For that purpose, we analyzed read and spontaneous data separately. As a result, even a superficial comparison revealed some peculiarities between the two given speech varieties.

First of all, read speech in general could be characterized by a greater number of "complex" pitch accents, i.e. bitonal and tritonal ones. Vice versa, monotonal pitch targets prevailed in spontaneous realizations. Thus, the frequent pitch accents L*H, L*HL and H*L were more common for the read speech samples, whereas H* and L* were more dominant in spontaneous speech.

More detailed analysis of each of the more frequent pitch accents brought with it some clear distributional preferences in the six aptitude groups under analysis in read and spontaneous speech.

Rising tones L*H

L*H accents, the most frequent category in read speech, were found on ~36-40% of all prominent syllables. The highest-aptitude speakers exhibited a clearly different frequency of occurrence pattern in this category: they only applied the rising targets to about a third of all stressed syllables in their realizations.

In spontaneous speech, the general percentage of L*H accents was lower (~22-28%), whereby the two higher-aptitude groups, as well as the low-aptitude informants, tended for lower values than the remaining subjects.

Simple high targets H*

The simple high targets H* were the most frequent category in spontaneous samples and appeared to be the second most frequent intonation event in read speech. In the latter speech variety the greatest amount these pitch accents could be observed in the below-average group (~33%) and the smallest – as realized by the above-average speakers (~25%). For highest-aptitude and low-aptitude subjects the distributional pattern was almost equal (with ~29%).

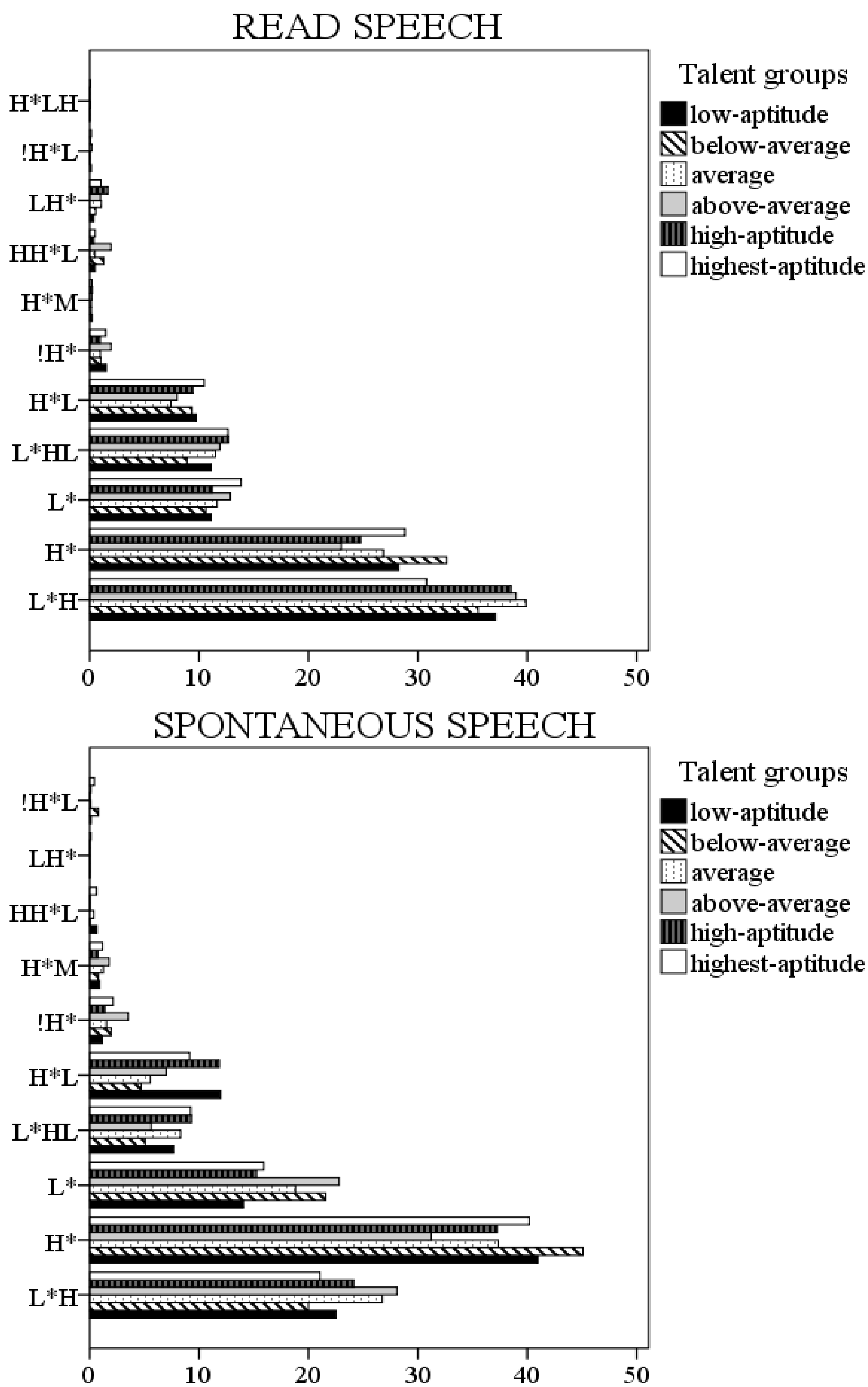


Figure 6.3: Distribution of pitch accents in all aptitude groups by speech type – read versus spontaneous realizations in the German-language part of the experimental corpus . On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

Still a slightly sparser employment of H* accents was demonstrated by the average and high-aptitude test persons (~25-27%).

In spontaneous speech, we observed a similar cross-group distribution pattern as in the read speech, only with a larger general percentage of H* events. Thus, H* accents were also most common for the below-average group (~45%) and least characteristic of the above-average speakers (~32%). As in the read speech realizations, highest-aptitude and low-aptitude speakers produced simple high targets with an almost equal frequency of occurrence (~40%). The only difference in the distribution of H* pitch accents between read and spontaneous samples consisted in that fact that in the latter speech type high-aptitude subjects replicated the average-group pattern in this category, applying high tone targets to ~38% of all prominent syllables. As mentioned above, the percentage of the pitch accents in question in read speech was lower in this group, compared to average speakers.

Simple low targets L*

The third most frequent category in spontaneous speech was the L* pitch accent, accounting for about a fifth of prominent syllables. More specific distribution of the given events in the aptitude groups was as follows: the above-average, average and below-average groups had the greatest amount of L* targets in their realizations (~19-24%); the remaining informants had a slightly lower percentage of simple low targets (~14-16%).

The distribution picture in the talent groups under analysis in read samples was rather homogeneous: the category in question accounted for ~11-14% of prominent syllables in all groups.

Rising-falling tones L*HL

The frequency of occurrence of rising-falling L*HL accents in read realizations, rise-falls were almost equally frequent as the simple low targets and presented a relatively regular distribution pattern of slightly decreasing numbers of rise-falls in most groups – accounting for ~11-13% of all prominent syllables. The lowest count of rising-falling targets was observed in the below-average group (~8%). In spontaneous speech below-average and above-average speakers employed the smallest amount of L*HL accents (~6%). The other groups realized a comparable number of prominent syllables with the rising-falling targets, with percentages being slightly higher for the highest- and high-aptitude groups (~10%) and marginally lower for the average and low-aptitude subjects (~8%).

Falling tones H*L

The final frequent category, i.e. the falling H*L accents, accounted for roughly the same amount of prominent syllables in read and spontaneous speech. However, as was true for the other frequent events, there was more homogeneity between the talent groups in read speech than spontaneous speech realizations. Thus, in the former speech type the percentage of falling tones in all groups did not exceed ~8-10%.

Spontaneous speech samples revealed marginally higher percentages of H*L accents in utterances rendered by most groups (~8-12%), except for the below-average group with a lower frequency of occurrence of the given events (~5%).

Figure 6.3 is a detailed illustration of the trends detected for all aptitude groups in German by speech type.

At the final stage of pitch accent analysis we split the data by applying two factors – speech type and gender, in order to see, to which extent the more general corpus trends have been influenced by them.

Corpus analysis by speech type and gender

On a general note, it should be mentioned that the general tendencies revealed in previous examinations, i.e. when analyzing the whole corpus in general as well as with the application of speech variety and gender factors, were present in this detailed investigation as well. Namely, five frequent categories were found in all talent / speech type / gender samples – L*H, H*, L*, L*HL and H*L. Moreover, their distribution in gender / talent groups in read and spontaneous speech also supported the general tendency: of the two most frequent categories, L*H was employed more frequently in read utterances, whereas H* was more common for spontaneous speech realizations. Read speech samples, especially realized by female speakers, exhibited fewer mismatches between the talent groups than spontaneous utterances.

Regarding the gender peculiarities between the talent groups in the two speech types, we found quite a few differences between male and female speakers in that respect.

Highest-aptitude group

For example, the visibly smaller number of L*H accents in the highest-aptitude group, which we mentioned earlier, appeared to be more characteristic of female speakers, and especially in the read speech utterances (~42% vs. ~27-40% in the remaining groups in this condition).

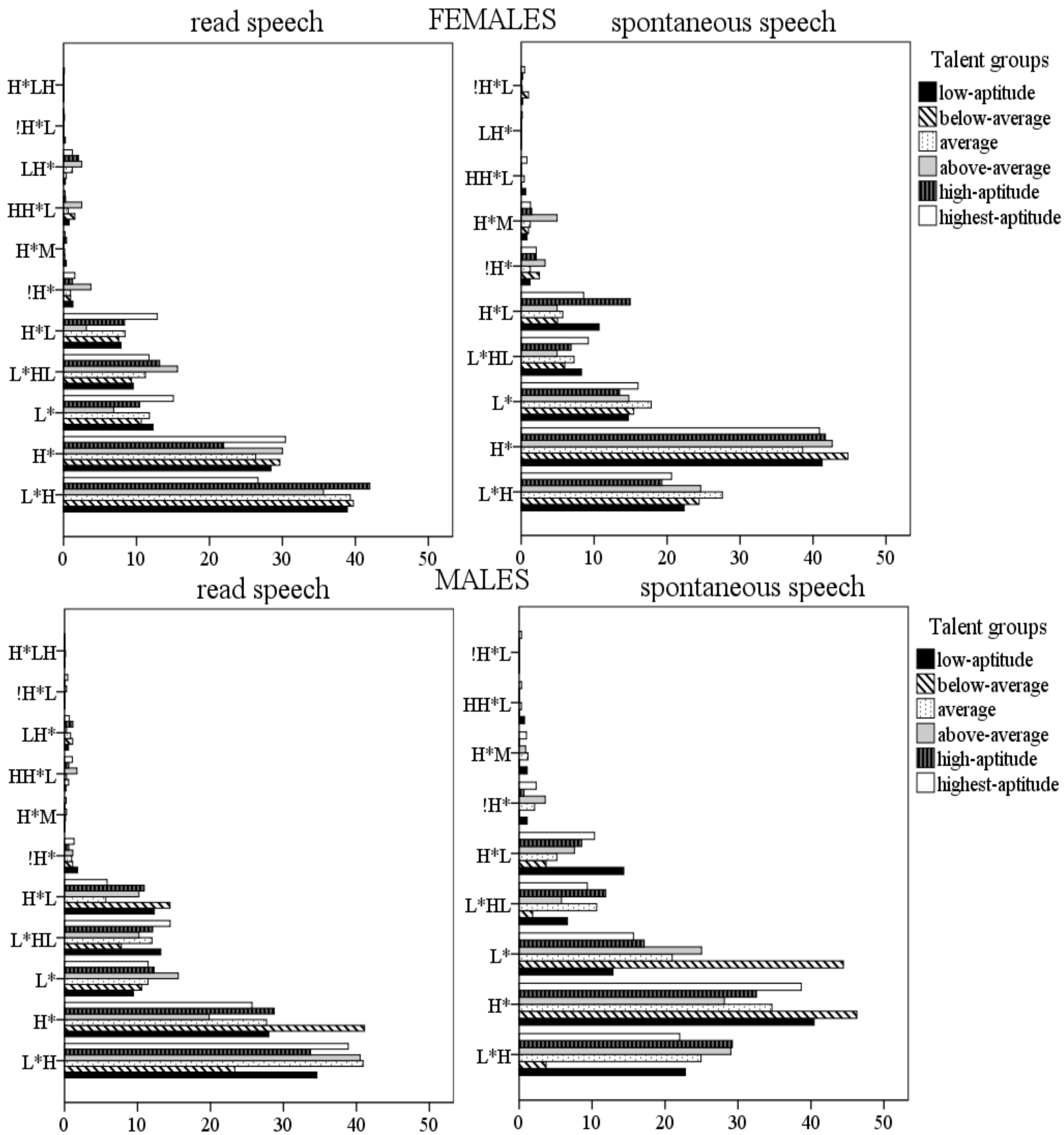


Figure 6.4: Distribution of pitch accents in all aptitude groups by gender – male versus female realizations – in read and spontaneous speech in the German-language part of the experimental corpus. On the x-axes – frequencies of occurrence within a group in per cent; on the y-axes – the observed pitch accents

Male speakers of highest aptitude used the rising tones on ~40% of all prominent syllables in read speech, as opposed to only ~27% employed by the female subjects. In spontaneous speech male and female highest-aptitude test persons exhibited a similar distribution pattern of L*H events, which accounted for around a fifth of all investigated syllables.

In the other frequent ToBI categories, the highest-aptitude speakers of both genders generally demonstrated comparable frequencies of occurrence. More specifically, the simple high targets H* accounted for ~40% of all accented syllables in spontaneous speech and ~27-30% - in read utterances, whereby the male subjects had a slightly lower number of the tokens in question than the females in the latter condition. The number of L* events in spontaneously produced utterances in both genders was ~15%. In read speech, however, we found a mismatch with respect to the frequency of occurrence: the number of the low-target tokens produced by the female speakers was ~15%, or ~5% higher than for their counterpart male group. Analyzing the rising falling tones, we could see that these events were common for about a tenth of all the accented syllables in the highest-aptitude group corpus, except for the performance of the female speakers in read speech, were this percentage was increased by around 4%.

Finally, one of the greatest discrepancies between the genders could be found within the H*L category in read speech, where female subjects produced about twice as many tokens with these intonation events than males (~13% against ~6%).

High-aptitude group

There was less agreement between the high-aptitude male and female speakers, as far as the distribution of ToBI categories is concerned. In read speech, for instance, the most frequent category for all subjects in this aptitude gradation was the L*H accent, followed by the H* events. However, the ratio between these two categories was different for the informants of high pronunciation aptitude gender-specifically: female speakers took a consistent preference of the rising tones (~42%) over the simple high targets (~22%); for the male subjects the mismatch between L*H and H* accentswas less pronounced (~34% vs. ~29%, respectively).

In spontaneous speech samples the distribution pattern was somewhat reverse. Here, the H* targets were the most frequently used category overall. As in read speech, the percentage difference between L*H and H* events was much more prominent within the female test persons (~20% vs. ~42%, for the two categories, respectively), as opposed to male speakers (~29% vs. ~33%).

The rest of the frequent categories in read speech indicated comparable numbers of tokens in them across the two genders: ~11-13% for L* accents and ~9-11% for the H*L targets.

The distribution of these three ToBI categories in spontaneous speech was not as homogeneous. Thus, the high-aptitude female speakers more extensively employed H*L accents (~15%) than the male test persons (~9%). Finally, the simple low targets were more frequently found in the male productions (~18% vs. ~14% in female speech).

Above-average group

The above-average speakers constituted one of the smaller groups, whereby there was only one female subject classified within this aptitude gradation. Despite this fact, her performance appeared to comply with the general distributional pattern. The two male speakers, on the other hand, demonstrated deviating performance in several conditions.

When examining read speech, we could state that the rising tones L*H accounted for ~35% of all the accentuated syllables in female productions, which complied with the corresponding percentages in other female groups. For the male subjects of above-average pronunciation aptitude the given category was more frequent, applied about 40% of all tokens, which was also the highest percentage in this condition of all male groups – a performance only replicated by the average male subjects.

With respect to the H* accents, we found that the female above-average informant followed the general trend discovered for the other groups by employing the given category in ~30% of all prominent syllables in read speech. Their male counterpart group, on the other hand, produced a smaller number of H* tokens (~25%).

Both female and male above-average speakers deviated from the other groups in the realization of L* pitch accents: the male subjects applied them more extensively (~17% vs. ~13% in the other male groups), whereby the female subject reduced the percentage of simple low targets as compared to the remaining female test persons (~8% vs. ~13%).

The percentage of L*HL accents in above-average male realizations was comparable to that in the other groups, comprising ~10% of all syllables. The female subject, on the contrary, employed rising-falling tones more extensively than the other female subjects (~15% vs. ~10-13%).

Also deviant was the behavior of the above-average female speaker with respect to the falling tones. They only made a small proportion of syllables (~4%), as opposed to the average number of ~10% in both male and female groups.

In spontaneous speech, we found more commonalities between male and female above-average speakers. More precisely, the number of L*HL tokens was almost equal as produced by both genders (~5%). The proportion of falling tones was also similar (~6-8%). Notably, these numbers also for the most part agreed with the ones spotted in the remaining groups.

As noted before, there was a general preference by most speakers for L*H accents in read speech and for the simple high targets H* in spontaneous speech. This trend could not be supported by above-average male data: the amount of both rising and H* accents was almost equal for these informants and accounted for around a third of the investigated syllables produced by them. The female subject's productions, however, did not fall out of the above-mentioned pattern, the number of H* events produced by this speaker reached ~40%, whereas the rising tones were applied to a fourth of the prominent syllables in spontaneous speech.

The final category under investigation, the L* pitch accents, was found to be more frequent for the male above-average speakers as contrasted to most other male groups (~25% vs. ~15-20%). The female subject could not be distinguished from the other female test persons in this regard by producing a similar number of simple low targets (~15%).

Average and low-aptitude groups

As opposed to the relatively unstable above-average group, average and low-aptitude test persons displayed an overall cross-gender agreement in the distribution of ToBI categories. We could only detect minor mismatches in percentages. The greatest discrepancy involved the low-aptitude male speakers, who produced more H*L targets than female subjects both in read speech (~12% vs. ~9%, respectively) and in spontaneous speech (~15% vs. ~11%, accordingly).

Apart from that, no outstanding observations could be drawn from the utterances pertaining to the average and low-aptitude groups, due to cross-gender agreement in these groups, as well as compliance with the general trends.

Below-average group

The major observation about the below-average subjects was the fact that the female part of this group could not be set apart from the other female speakers in performance – all their samples fully complied with the general frequency trends in both read and spontaneous speech.

The male below-average speaker only displayed non-deviant behavior in several cases, i.e. as far as the distribution of the less frequent categories (L*, L*HL, H*L) in read speech is concerned. In all the other conditions the distribution of the five basic pitch accents was clearly different.

Thus, below-average male subject produced fewer L*H tokens than the other male speakers. It was quite pronounced in read speech (~23% vs. over 30% in the other groups); in spontaneously produced utterances the difference was even more striking (under 5% vs. at least over 20% in the other groups). Also smaller were the percentages of H*L and L*HL events in spontaneous speech supplied by this subject (~5% vs. 7-15% for the former category; ~2% vs. 10-15% for the latter tokens).

The results of the detailed investigation of pitch accent distribution in the German part of the experimental corpus by gender and speech type are reflected in Figure 6.4.

The analysis of ToBI pitch accent distributions, provided in the current section, outlined the general tendencies in the aptitude groups, both generally and with regard to the gender and the speech type factors. We will further attempt to analyze the boundary tone categories in a similar fashion.

6.1.1.2 Boundary tones

General corpus analysis

The investigation of boundary tones in the German part of the experimental corpus revealed a lower general degree of cross-category variation due to two facts: firstly, the a priori smaller number of ToBI categories to describe these events, and, secondly, the nature of the produced utterances – mostly affirmative sentences.

As expected, the most frequent category on this level was the full low boundary L-L%, accounting for around 40% of all break markers in the corpus. The full high edge tone H-H% and the so-called default boundary % marked a substantial number of intonation phrases as well, each delineating just under a third of IP (intonation phrase) final pitch levels / configurations. To the frequent categories in the data also belonged the intermediate low L- and high H- boundary tones, which were common for about a tenth of ip (intermediate phrase) edges[15]. There were also some aptitude-group specific peculiarities in the distribution of these frequent boundary tone events.

Low full boundary tones L-L%

With respect to the L-L% tones, only the average speakers deviated from the general pattern by producing a smaller number of low full boundaries (~34%). The other groups followed the above-mentioned frequency of occurrence trend with ~37-40% of IP's ending on a low pitch.

15 Following Beckman and Pierrehumbert's definition (1986), we distinguish between a full intonation phrase (IP), demarcated by a clear pause in the speech signal, and an intermediate phrase (ip), which is perceived as a complete chunk in an utterance, but is not followed by a full-fledged pause.

High full boundary tones H-H%

The full high boundary tones H-H% enjoyed the highest frequency of occurrence in the above-average group, with ~35% of all break markers. Lowest was the number of tokens pertaining to the category in question, as realized by the below-average speakers (~25%). For the rest of the subjects, about a third of the intonation phrases ended at a high pitch level.

Default boundaries %

A clear mismatch between the groups could be found within the category of default boundary. Here, the average, below-average and low-aptitude speakers demonstrated the highest frequency values (~26%). The number of default boundary tokens was visibly lower in the high-aptitude and highest-aptitude groups (~20%) and lowest – in the above-average speakers' realizations (~17%).

Intermediate high H- and low L- boundary tones

In the final two frequent categories, i.e. the intermediate low L- and high boundaries H-, the distribution patterns were for the most part homogeneous – with ~4-8% of boundaries characterized by a low or high intermediate edge token. The groups producing marginally smaller number of these ip events (by ~2%) were below-average and low-aptitude subjects for L-category and the low-aptitude subjects for the H- tones.

Figure 6.5 is a detailed illustration of the distribution of boundary tones in the whole German-language corpus.

Following the same logic as in the investigation of pitch accents, we further looked into the possible gender differences in the distribution of boundary tones.

Gender-specific analysis

The first conclusion we could draw from analyzing boundary tone distribution in the German corpus by gender was the fact that both female and male samples for the most part agreed with the general distributional pattern described for the entire German database, with some deviations in each of the frequent boundary tone types.

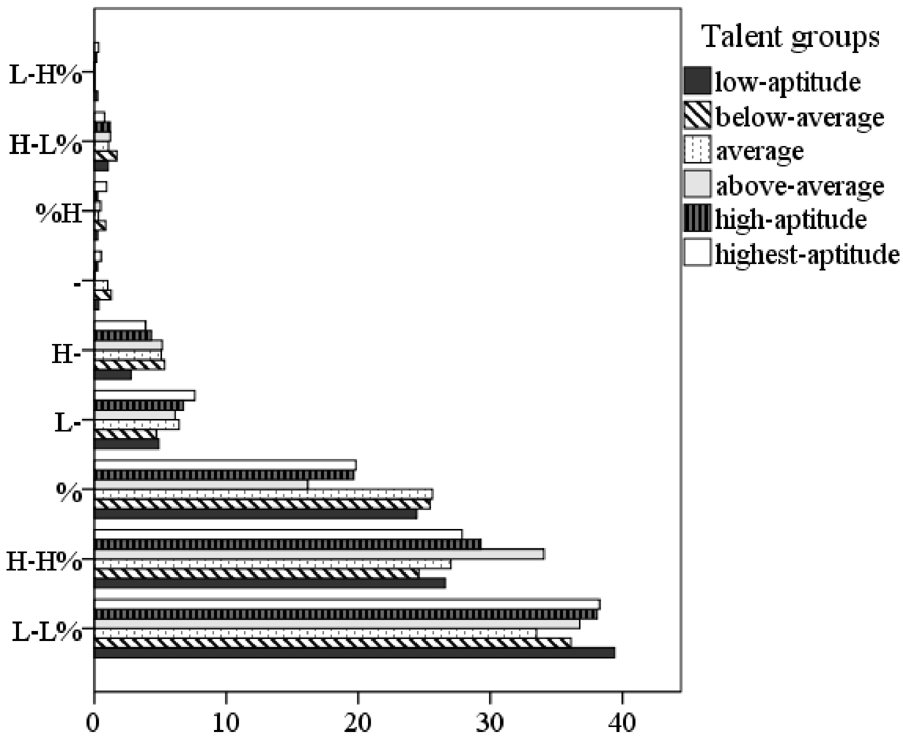


Figure 6.5: Distribution of boundary tones in all aptitude groups in the German-language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

Low full boundary tones L-L%

Thus, we found almost equal percentages of L-L% events in most counterpart gender groups, with male / female mismatches not exceeding 2-3% and the general number of low full boundaries accounting for ~34-40% of all IP's. The low-aptitude and especially above-average groups, on the other hand, presented the biggest deviation between male and female test persons in the current category. For the former group the visible, but not drastic, difference consisted in the following ratio: ~42% of L-L% tones in male speech against ~38% in the female realizations. Above-average speakers differed even more conspicuously: there was a 10% divergence of the female realizations from the male speech (~45% vs. ~35%, accordingly).

High full boundary tones H-H%

In the full high intonation phrase boundaries, only the speakers of highest pronunciation aptitude of both genders agreed as to the number of H-H% tokens (~32%). The remaining samples contained observable discrepancies genderwise, which, in turn, resulted in deviations from the general-corpus figures.

More specifically, average, below-average and low-aptitude female informants had ~3-5% more H-H% events in their speech than the counterpart male groups (the general percentage fluctuating between ~24% and ~29%). In the high-aptitude group, the picture was reverse: male speakers produced the full high boundaries more extensively than the females (~33% vs. ~28%).

Still the greatest mismatch in the category in question was registered for the above-average subjects: the male subjects had the greatest proportion of high full boundary tones of all the investigated L1 German groups (~38%), whereby the percentage rendered by the female test person of above-average ability was ~15% lower.

Finally, the below-average male subject was singled out for the lowest frequency of occurrence of full high boundaries of all male and female groups (~20%).

Default boundaries %

The default boundary category was marked by a number of consistencies in male and female realizations as well as some differences in this respect.

To the cross-gender similarities belonged the performance of the highest-aptitude (~20%), above-average (~15%) and below-average (~25%) speakers. The remaining three groups presented examples of dissimilar distribution patterns across the genders. The high-aptitude as well as low-aptitude male productions were characterized by ~5% decrease of default boundary events as contrasted with the corresponding female samples (~27% vs. ~22% for the latter group; ~22% vs. ~17% for the former). The difference between the average speakers took a reverse direction, with male speakers realizing more tokens with the % boundary tones than the females (~30% vs. ~23%).

Intermediate high H- and low L- boundary tones

The intermediate high boundary tone category appeared to present a relatively homogeneous picture from the cross-gender perspective – the mismatches between the groups were only slight, as these events made ~5-8% in the corpus in all aptitude groups.

In the L- events the cross-gender discrepancies were generally more pronounced. The greatest one involved the above-average test persons. The female subject realized a visibly greater number of intermediate phrases at a low pitch level of all groups in general and the male above-average speakers in particular (~12% vs. ~5%). Other cross-gender differences included the below-average and low-aptitude informants, for whom the number of L- tokens appeared to be doubled in male speech as opposed to the female productions (~4% vs. ~8% for both groups).

Figure 6.6 is to be referred to for the illustration of all findings in boundary tones by gender in the German-language part of the experimental corpus.

Next we compared the read and spontaneous speech realizations.

Corpus analysis by speech type

A general inspection of the boundary tone distribution by speech type allowed us to draw a few general conclusions about read and spontaneously produced utterances in all of the aptitude groups.

Read speech yielded more predictable frequency of occurrence / distribution patterns. Thus, there was a clear preference for the low full boundary, which can most probably be accounted for by the fact that the two read-speech abstracts, “The North Wind and the Sun” and the abstract from “Mrs. McWilliams”, were pieces of narration. The proportion of the other frequent categories, in turn, was conspicuously lower.

As noted previously for pitch accents, there was also less overall stability and regularity in boundary tones in spontaneous speech.

Taking due account of the general trends mentioned above for read and spontaneous speech, it ought to be mentioned that there were group-specific peculiarities, which made the samples deviate from the overall patterns to some degree.

Low full boundary tones L-L%

Thus, the low full boundary L-L%, the most frequent category in read speech, was most common for ~40-45% of all IP’s, as realized by all aptitude groups. In spontaneous speech the pattern was slightly different: whereby most groups applied a low full boundary tone token to ~30-35% of all phrases; only about a fourth of all intonation phrases had a full low boundary in the realization of the below-average, above-average, as well as average speakers.

High full boundary tones H-H%

The high full boundary was applied to ~24-28% of all intonation phrases in read speech, being most extensively used by the highest-aptitude as well as above-average groups and being least common in below-average subjects’ productions.

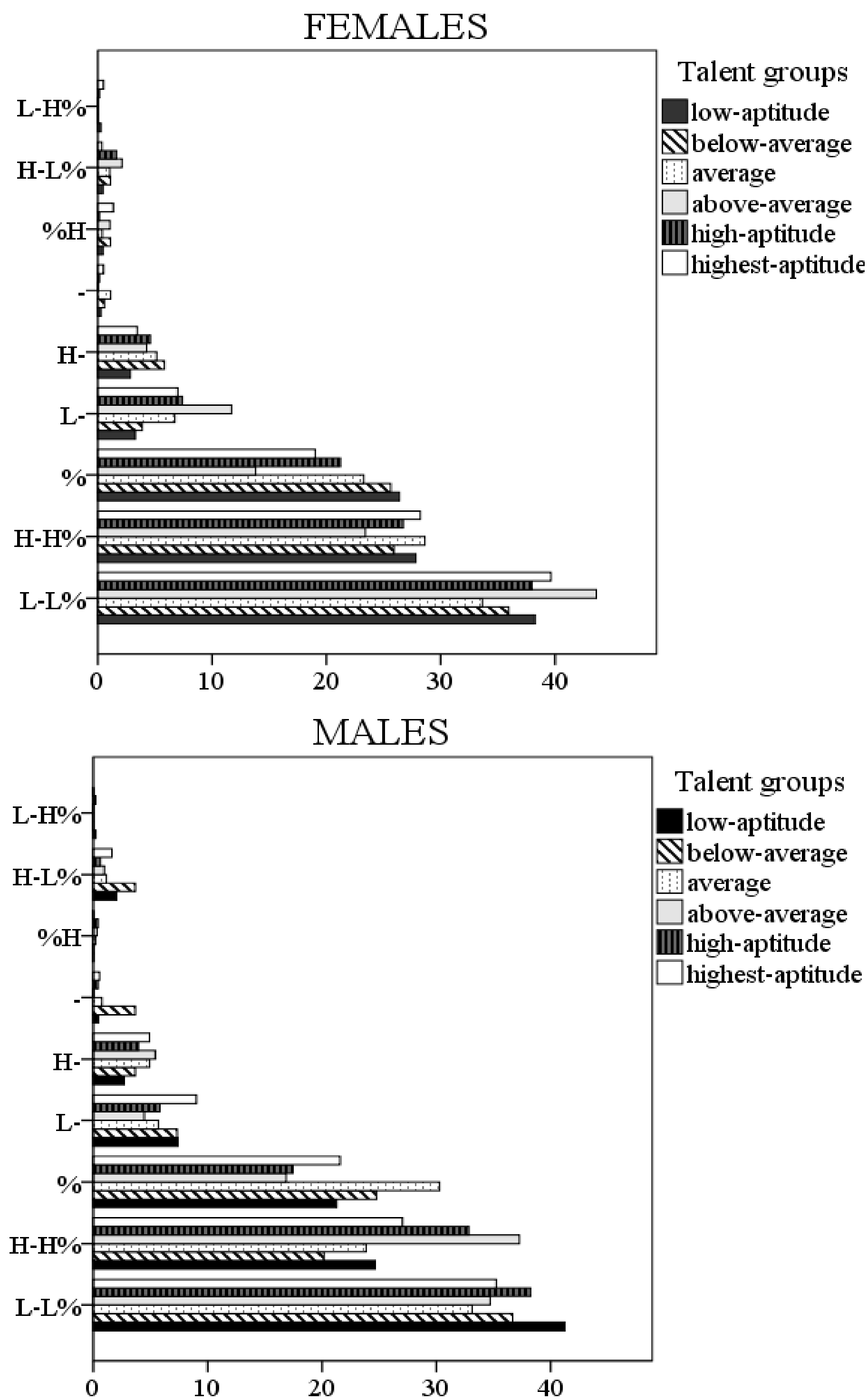


Figure 6.6: Distribution of boundary tones in all aptitude groups by gender – female versus male realizations in the German-language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

Spontaneous speech revealed a less homogeneous outline in the distribution of full high boundary events, where above-average subjects performed in the most idiosyncratic way, employing this edge tone type to about 40% of all intonation phrases. The corresponding percentages in the remaining groups were much lower (~25-30%).

Default boundaries %

The most peculiar distribution pattern could be found in the default boundary category % in read speech. Here, the figures split in half: low-aptitude, below-average and average speakers produced about 20% of all intonation phrases with this edge tone type; the percentages for highest-aptitude and high-aptitude, as well as above-average speakers were much lower (~10-12%).

In spontaneous speech the pattern was not so clear-cut. Below-average subjects employed this category more often than the other groups (~38% vs. ~25-32%, accordingly). Notably, the above-average produced the smallest number of such tokens (~25%).

Intermediate high H- and low L- boundary tones

The less frequent boundary tone categories, i.e. the high and low intermediate edge tone events, did not demonstrate any stable trends as far as their frequency of occurrence is concerned.

With respect to the H- events in read speech, the cross-group differences were only slight, as the given category accounted for ~3-9% of all ip's. In spontaneous realizations, however, it was possible to single out the low-aptitude and average groups with lower than average percentages of high intermediate boundary tones (~5% vs. ~10%, accordingly.).

The low intermediate boundaries were only common for ~5% of all ip's in the German-language corpus, and no striking cross-group differences could thus be discovered there.

Figure 6.7 reflects all the findings in the experimental data split by speech type.

The analysis of boundary tone distributions was finalized by applying the two factors, i.e. gender and speech type, simultaneously.

Corpus analysis by speech type and gender

First of all, it should be noted that the application of these two factors revealed a substantial amount of heterogeneity in the samples supplied by most aptitude groups, both dependent on gender and with regard to speech type.

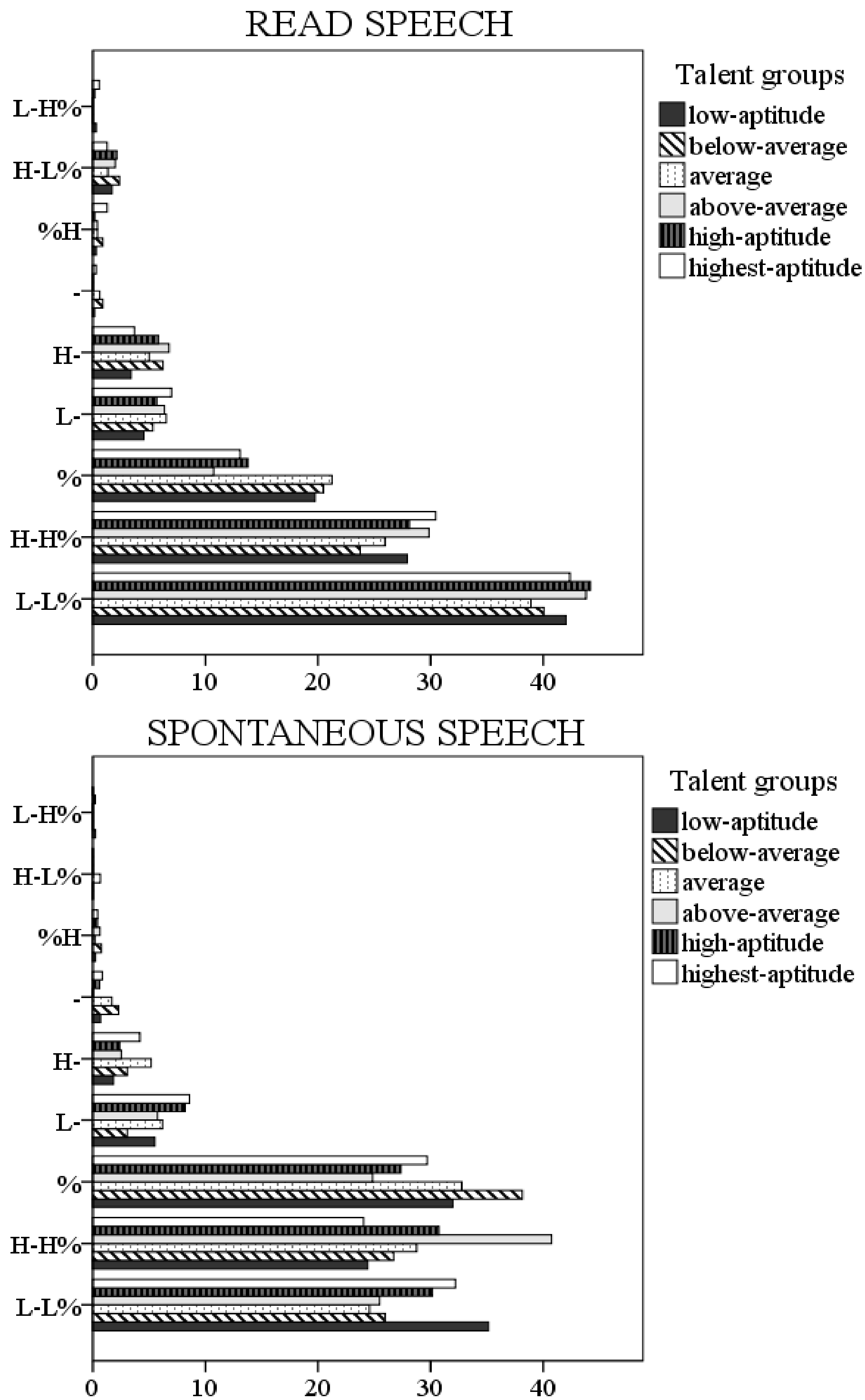


Figure 6.7: Distribution of boundary tones in all aptitude groups by speech type – read versus spontaneous realizations in the German-language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

Highest-aptitude group

Nevertheless the highest-aptitude speakers displayed considerable stability in this respect. The frequency of occurrence of most boundary tones was relatively equal as realized by male and female subjects, with only the exception of the default boundary category in read speech, where male speakers applied it to a substantially greater number of intonation phrases (~19%) than the female subjects (~11%). Still another, but less striking, difference concerned the distribution of low intermediate edge tones L-, which were more proper to the male part of the group (~11%) than the female subjects (~8%).

Other than these instances, the proportions of boundary tones were in accord with the figures we presented in the general-corpus, gender-specific and speech-type analyses.

High-aptitude group

High-aptitude male and female speakers also demonstrated a certain level of agreement in the realization of phrase boundaries. The greatest mismatches included the frequency of occurrence of the low full boundary in read speech, which was applied more extensively by male subjects (~49%) as opposed to female speakers (~41%). In spontaneous speech it was the distribution of full high boundaries H-H% that revealed a significant discrepancy in productions: male speakers realized ~40% of all intonation phrases on a high pitch level; H-H% events were only found in about a fourth of all female realizations. Contrary to that, there was a clear preference for the default boundary type by high-aptitude female test persons (~32%), which contradicted with the male-speaker pattern (~20%).

Above-average and below-average groups

The smallest groups, i.e. the above-average and below-average subjects, who often showed less consistency in the distribution of pitch accents, exhibited more regularity on the boundary tone level in the sense that there were fewer cases of these groups clearly deviating from the average percentages per category observed for the other speakers. Nevertheless the cross-gender differences within the above-average and below-average groups were quite apparent.

For the above-average subjects, the differences between males and females speakers could be detected in almost all categories in the two speech types. As far as read utterances are concerned, the most conspicuous discrepancy was found within the L-L%, where the female informant realized about 55% of all intonation phrases with a full low boundary token; the corresponding value produced by the male group, as well as the remaining male and female subjects, was ~15% lower. In a reverse fashion, the male test persons had more instances of the high full boundaries in their read speech (~33%) than the females (~21%). Also greater was the percentage of default boundary % events in male realizations (~12% vs. ~7% in the IP's pertaining to the female speakers). Finally, the mismatches in the less frequent categories L- and H- were not that striking.

Male and female speakers of above-average pronunciation aptitude also demonstrated unequal cross-gender performance in spontaneous speech, which was evident in all of the frequent boundary tone categories. Thus, there were fewer L-L% tokens in female speech as opposed to the male realizations (~15% vs. ~28%, correspondingly). The male subjects also exceeded females with regard to the H-H% category. In fact, they not only exhibited a greater percentage of such events compared to their female counterpart group (~43% vs. ~29%), this proportion of high full boundary tones was highest of all subjects in general, regardless of the gender factor. The above-average female informant, on the other hand, displayed a greater number of default edge tones in her speech (~30% vs. ~23% in male productions). The most noticeable cross-gender difference, however, could be seen in the L- category: the female subject had the greatest number of ip's finalized with the given tone (~20%), whereas in the male realizations L- tones were marked by a rather low frequency of occurrence (~3%).

For the below-average group, only one clear inter-gender difference could be found: in spontaneous speech the male speaker once again exposed idiosyncratic behavior by applying the default boundary to ~45% intonation phrases, whereas in female speaker realizations this category was present in ~35% of boundary markers.

Average group

There were two major differences between the male and female average speakers, and both pertained to the default boundary category, which was more frequent in the male productions in both speech types: the numbers were higher than the respective figures for female speakers by about 10% in each case, accounting for ~28% of IP's in read speech and ~38% in spontaneously produced utterances.

Low-aptitude group

Finally, male and female frequency of occurrence patterns in the low-aptitude group were marked by a considerable degree of homogeneity. Here, like with the average speakers, the main difference was found in the default boundary % category in read speech, where the female subjects produced about a fourth of all intonation phrases with these intonation events; for male speakers this number was about 10% lower. There was also a 5% difference in the distribution of low full boundaries L-L% in read utterances, which were more typical of the male subjects. Apart from that, in both speech types, low intermediate edge type L- was more common for low-aptitude male subjects (~7-8% vs. ~3% in all conditions). The remaining differences in the low-aptitude group were but marginal.

Figure 6.8 is to be referred to for the illustration of frequency of occurrence tendencies in boundary tones, influenced by the speech type and gender factors.

In the current section we summarized the distribution patterns of ToBI pitch accents and boundary tones in the six aptitude groups in their native language German. The data were analyzed, first of all, in the whole corpus in general and, following that, under the influence of two major factors – speech type and gender – both applied separately and in combination.

The tendencies observed for the subjects' L1 can be further compared to their performance in English, in order to track the instances of SLA phenomena we discussed in Chapter 2 or to state the absence of any cross-language interference. Hereby one of the most crucial objectives is to find out whether L1 / L2 processes had any systematic relation to a speaker's level of pronunciation aptitude.

6.1.2 English language data

6.1.2.1 Pitch accents

The second stage of the phonological investigation of ToBI categories consisted in a detailed analysis of the English part of the corpus, the target of the current investigation, whereby the German data described in Section 6.1.1.1 should be taken as the core to any comparison and trend identification.

On the whole, the same categories were identified as frequent in the English data as the ones we singled out for German. Thus, the simple high targets H* and the rising tones L*H had the highest frequency of occurrence in all groups. The distribution of the falling targets H*L, simple low tones L* and the rising-falling accents L*HL was more scarce, but these categories could still be considered to be relatively frequent in the corpus. We will therefore focus our analysis on these five frequent pitch accents.

The remaining ToBI pitch accents were only common for a very low percentage of accented syllables and did not demonstrate any systematic trends with regard to their observed frequency of occurrence. Thus, they will not be included in the current investigation.

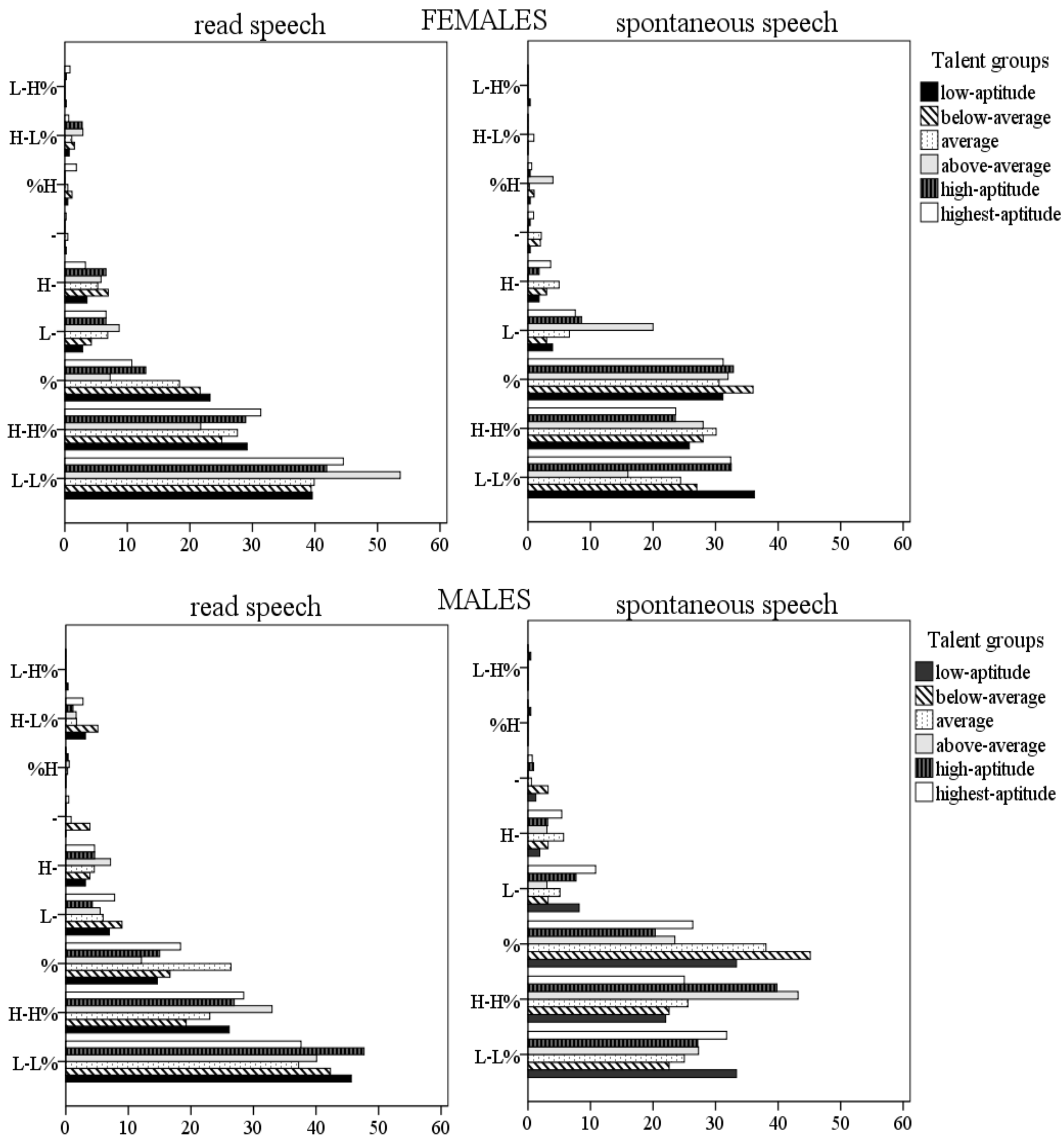


Figure 6.8: Distribution of boundary tones in all aptitude groups by gender – male versus female realizations – in read and spontaneous speech in the German-language part of the experimental corpus. On the x-axes – frequencies of occurrence within a group in per cent; on the y-axes – the observed boundary tones

General corpus analysis

The initial glimpse at the figures in the distribution of pitch accents in English revealed both features dependent on a speaker's native language, i.e. proper to L1 English and L1 German subjects, and aptitude-specific phenomena.

Simple high targets H*

In the most frequent pitch accent category, the simple high targets H*, native speakers exhibited a clearly different distribution pattern than that produced by all the German groups. More specifically, the frequency of occurrence of this category in the native group reached ~57% of all prominent syllables. For all the German speakers, on the other hand, the number of H* tokens did not exceed ~38-42%, with this highest percentage being registered in the high-aptitude group.

Rising tones L*H

Conspicuous was the distribution of the rising accents L*H, which are very typical of German speech in all parts of an utterance, but are rather sparse in English. Our data revealed a quite peculiar fact in this respect: the highest-aptitude and high-aptitude speakers employed this category far less frequently than all the other German groups (~21-22% vs. ~30-32%, respectively). This figure was still higher than the one obtained for the native speakers (~11%). But this intermediate position of the highest-aptitude and high-aptitude group appears to reveal a tendency of complying with the English-language trends, especially if one compares a higher percentage of rising tones in their L1 speech (~30%). It is also worthy of mentioning in this respect that this finding supports our previous research on the matter (Anufryk 2009, Anufryk & Dogil 2009). The current investigation therefore once again showed that the top aptitude speakers actually tend towards the English pattern to a greater degree than all the other groups. Above-average subjects, though possessing a relatively high pronunciation aptitude level, clustered with these other groups.

Falling tones H*L

An interesting observation can also be drawn from the distribution of the falling H*L accents. Bearing in mind that the falling configuration is generally more typical of English, based on earlier studies (e.g. Gimson 1989 for English; e.g. Mayer 1995 for German), it appeared that the highest-aptitude speakers actually slightly enhanced this pattern, applying the H*L targets to more syllables than native speakers themselves (~18% vs. ~15%, respectively). Falling tones were found in ~9-15% of all investigated cases in the realization of the remaining groups.

Simple low targets L* and rising-falling tones L*HL

The simple low targets L* set the below-average and low-aptitude speakers slightly aside from the remaining subject pool: the given events were granted a smaller number of prominent syllables than by the other groups (~7% vs. ~10-12%, accordingly).

The final frequent category L*HL revealed no striking trends, accounting for ~4-7% of all prominent syllables in the English-language part of the experimental corpus.

Figure 6.9 presents a detailed illustration of pitch accent distribution in the English part of the experimental data.

An additional insight into the data was found when we looked into the gender peculiarities of the distribution of pitch accents in English.

Gender-specific analysis

Simple high targets H*

By applying the gender factor to the data, we could state that the male native speakers were responsible for the significantly higher overall proportion of H* events in the corpus by L1 English speakers in general: ~62% of all prominent syllables carried a simple high target in native male speech, whereby most of the L1 German male groups only applied this category to ~40-45% of tokens on average. There were, however, two exceptions to this pattern. The below-average male informant demonstrated an outlying frequency of occurrence of the H* category of all L1 German male groups (~50%), which was only exceeded by the native subjects mentioned above. Still another outlier in this condition was represented by the average speakers with the lowest number of H* tokens in their realizations (~32%).

The mismatch between native and non-native speakers was not as pronounced as in male speech. In fact, native female subjects demonstrated a similar percentage of high full boundaries as the high-aptitude group (~48% and 45%, respectively). The other female test persons exhibited smaller amounts of simple high targets in their speech (~35-40%).

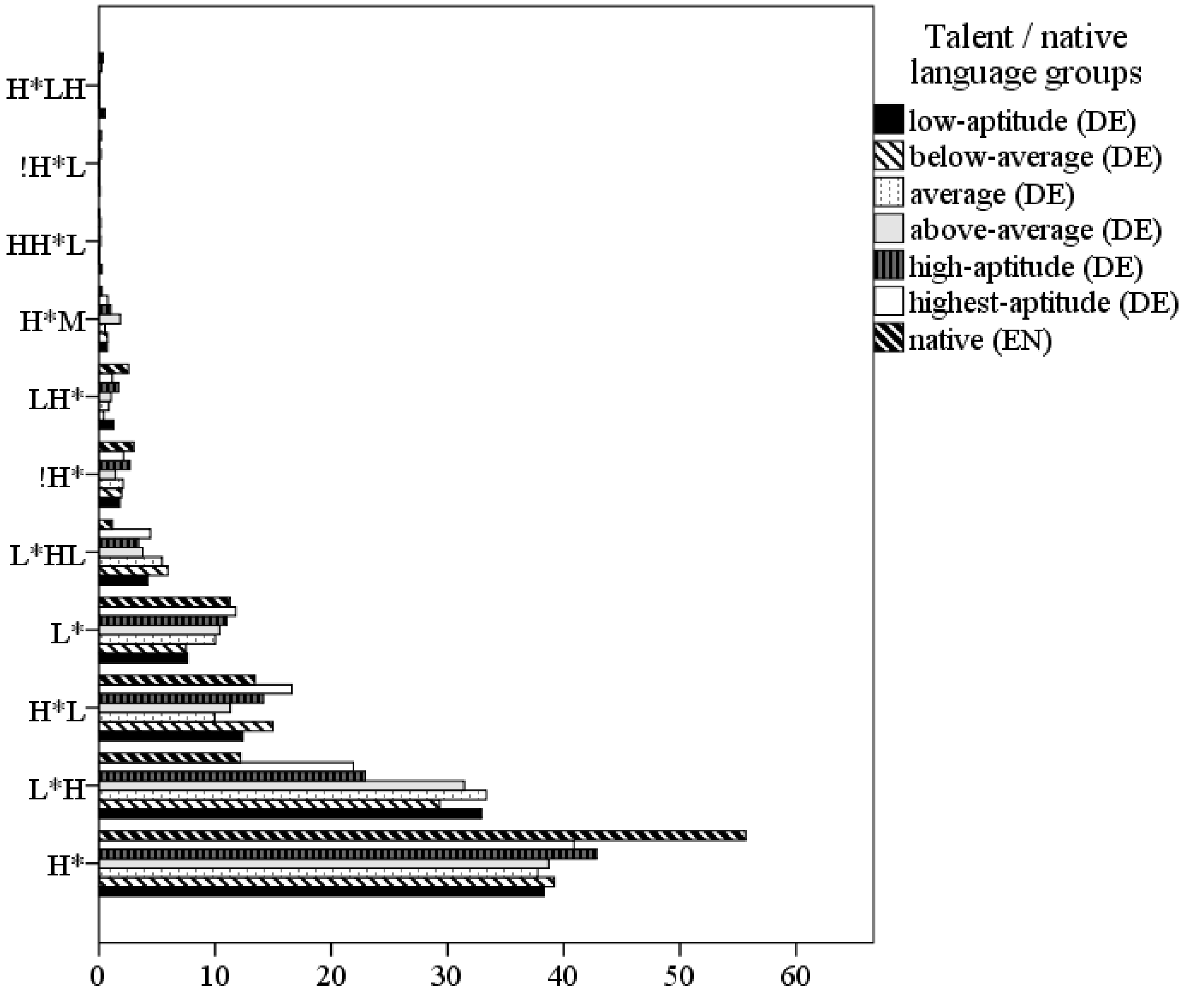


Figure 6.9: Distribution of pitch accents in all aptitude groups in the English language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

Rising tones L*H

When analyzing L*H accents genderwise, we could, s in the German part of the corpus, discover a frequency of occurrence pattern in female speech, concurrent with the aptitude gradation. It established gradually decreasing amounts of L*H events from the low-aptitude and below-average (~35-38%); average and above-average (~27-30%); highest-aptitude and high-aptitude (~21-22%); and, finally, native subjects (~18%).

The male subjects exhibited a less homogeneous pattern in this category. As was the case in several other instances earlier, the below-average male speaker performed in an idiosyncratic way in this condition, by applying the rising targets to only ~12% of all accented syllables, as opposed to a ~24% minimum in the other groups. The greatest number of L*H accents could be found in the average group (~40%), followed by the above-average (~34%). The highest-aptitude, high-aptitude and, unexpectedly, low-aptitude male speakers were closest of all L1 German groups to the native informants in performance, but their values were still much higher (~24-27% vs. ~8%).

Falling tones H*L

A clear frequency of occurrence pattern could be drawn from the distribution of H*L accents in female speech: there was a clear preference for this category by highest-aptitude and high-aptitude subjects as opposed to the other groups (~18% vs. ~10-11%, respectively).

In the male-speaker realizations, the below-average subject once again demonstrated idiosyncratic behaviour using the falling accent on about a fourth of all the prominent syllables against ~10-16% in the other groups.

Simple low targets L*

The simple low targets L* in the female realizations were most characteristic of the above-average, native and highest-aptitude speakers (~14-17%). For the rest of the groups this category was less frequent (~8-10%). In the male productions the simple low targets were applied to ~10-13% of all prominent syllables. The given events were marginally less characteristic of the low-aptitude and below-average (~6-8%) groups.

Rising-falling tones L*HL

No stable trends could be detected in the rising-falling L*HL category. The frequency of occurrence of the given events was comparable in most female and male groups, which accounted for ~5-8% of all prominent syllables.

Figure 6.10 is a detailed illustration of pitch accent distribution in the English-language part of the corpus genderwise.

When we applied the second influence factor to the data analysis, i.e. the subdivided the corpus into two speech types, we could not see any clear-cut differences between read and spontaneous speech, as we observed in the German part of the data. There were, however, some peculiarities proper to the aptitude groups, but no global trends.

Corpus analysis by speech type

Simple high targets H*

The simple high targets H* remained the most frequent category in both the speech types. As the general analysis suggested, this category was most characteristic of the native group (~54% in spontaneous speech and ~58% in read utterances).

In read speech realizations the native speakers were followed by the high-aptitude group (with ~44%), whereas the values in all the other German groups were almost equally lower (~38-39%).

The spontaneous realizations revealed a different distribution pattern: highest-aptitude, below-average and low-aptitude speakers employed the simple high targets similarly – on ~47% of all prominent syllables; the remaining groups used the given ToBI events more sparingly, applying them to ~40% of all tokens under investigation.

Rising tones L*H

The general tendency in the distribution of the rising L*H accents was but partially confirmed across the two speech types. The overall pattern, as the one revealed in the whole corpus, was almost replicated in the read speech realizations, i.e. the highest percentages of the rising targets were found in the low-aptitude, below-average and above-average, as well as the average group (~30-32%). In highest-aptitude and high-aptitude productions the rising accents were applied to about a fifth of the prominent syllables, which did appear to approach the native-speaker values (~9%) to the greatest degree.

In spontaneous speech only the highest-aptitude and below-average subjects demonstrated similar performance to the native informants, with 20-23% of rising tones in their speech. The remaining group samples contained about a third of L*H targets in them.

Falling tones H*L

No definite pattern could be detected in the falling H*L accents, especially in spontaneous speech. Here, the frequency of occurrence of the given events fluctuated between 9-12% in most groups, and a slightly higher percentage demonstrated by the highest-aptitude speakers (~15%).

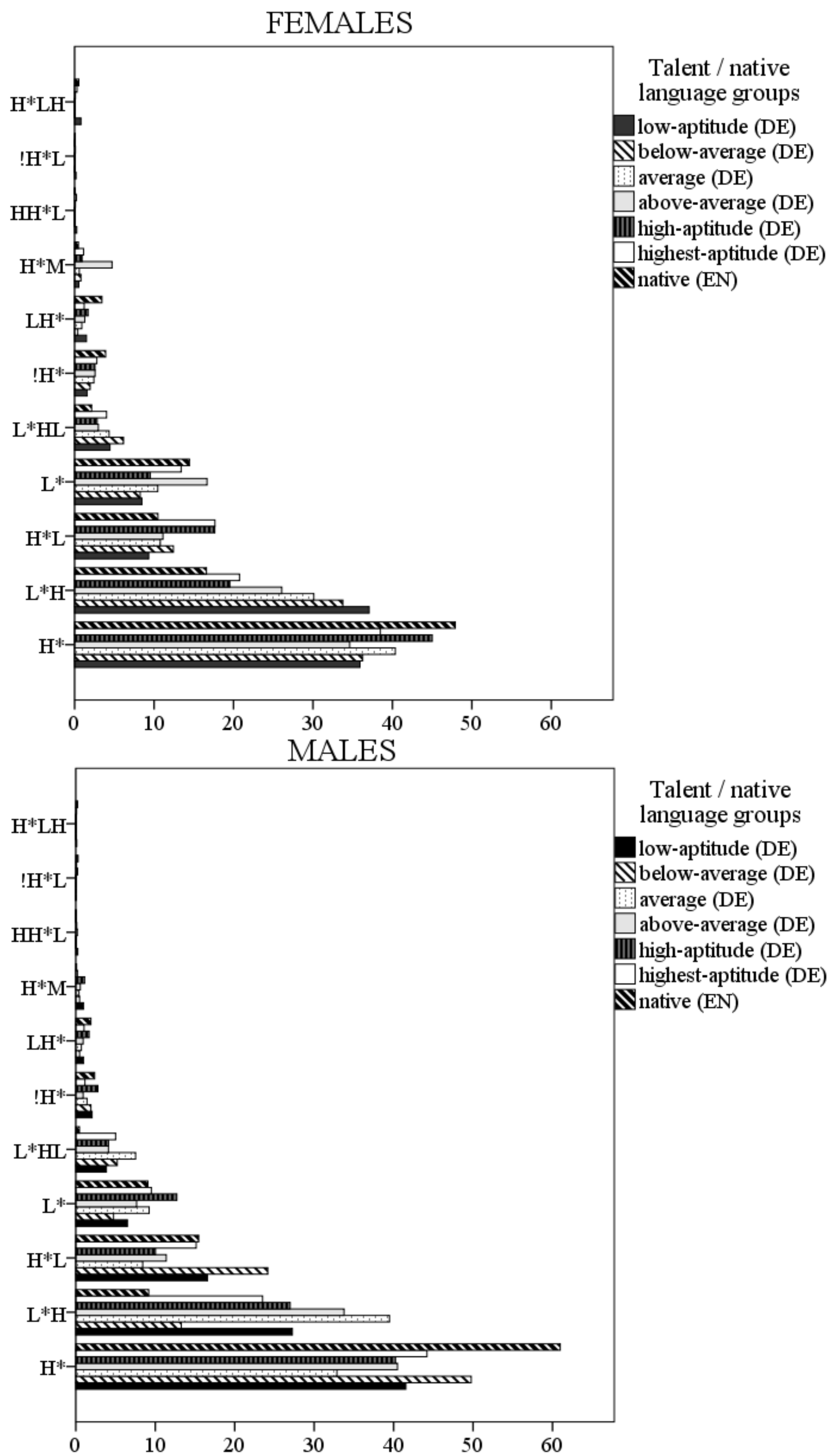


Figure 6.10: Distribution of pitch accents in all aptitude groups by gender – female versus male realizations in the English language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

In read speech, cross-group differences were also only marginal: the proportion of falling tones did not exceed ~12-17% in all groups.

Simple low targets L*

The simple low targets revealed a slight aptitude-specific tendency in read speech: native speakers, as well as subjects of higher pronunciation aptitude levels, i.e. highest-aptitude, high-aptitude and above- average, exhibited a slightly greater preference of these intonation events (~12-13%) than the average, below-average and low-aptitude speakers (~6-8%).

The low-aptitude group also demonstrated the lowest percentage of simple low targets in spontaneous speech (~7%). In the remaining groups the number of L* targets was greater: about 10-15%.

Rising-falling tones L*HL

One of the major observations to be made for the rising-falling tones was the fact that these events could not be considered a frequent category in spontaneous speech. Rather, their frequency of occurrence lay by ~3%, as in the infrequent categories.

In read speech, the corresponding percentages were higher - up to ~6-8%.

Figure 6.11 is to be referred to for the illustration of pitch accent distribution in the English part of the corpus by speech type.

Finally, we applied both the influence factors to the pitch accent data – those of both speech type and gender.

Corpus analysis by speech type and gender

Highest-aptitude group

The speakers of the highest pronunciation aptitude demonstrated stable performance genderwise and across the speech varieties, which also agreed with the general distribution patterns described for the entire corpus, gender-specific data and across the speech types. There were no striking differences between the male and female speakers overall, except for the L* accents in read speech, which were more common for the female speakers than they were for the male subjects (~7% vs. ~4%, correspondingly).

High-aptitude group

Several discrepancies were discovered between male and female speech, when we looked at the high-aptitude group realizations.

In the case with the L*H events, spontaneous speech revealed a cross-gender peculiarity with respect to their frequency of occurrence. Thus, the number of rising tones was visibly higher in male productions than in the utterances rendered by the female informants (~38% vs. ~22%, respectively). The falling H*L tones were generally more common for female test persons in both speech types (~20% vs. ~16% in read speech; ~12% vs. ~4% in spontaneous speech).

The remaining differences between male and female high-aptitude subjects were only marginal.

Below-average and above-average groups

Of the smallest groups, i.e. below-average and above-average speakers, the former group certainly exhibited less regularity, whereas the latter produced fairly regular patterns across the two genders.

The most prominent difference between male and female speakers of above-average pronunciation aptitude in read speech could be discovered within the simple low targets L*, where the female subjects produced more than twice as many tokens in question (~19%) than the male test persons (~8%). Apart from that, the frequency of occurrence of the rising tones was also uneven across the genders (~37% in male productions vs. ~26% in female realizations).

In spontaneous speech the major mismatch was detected in the H* pitch accents: the female part of the group applied simple high targets to only a fifth of all prominent syllables; the male speakers, on the other hand, exhibited values close to those of the other male groups (~44%). Still another striking dissimilarity of above-average female subjects as compared to all the other groups was observed in the infrequent categories, namely, H*M accents, which were employed in about a fifth of all prominent syllables by these speakers.

The major difference of female below-average speakers from the male subject of the corresponding pronunciation aptitude level was the fact that the female subjects fully complied with the general distribution patterns, whereas the male informant departed from them in quite a few instances throughout the current investigation, the gender / speech type analysis in English being no exception. It was seen most clearly in the H* category, where the below-average male speaker applied the simple high targets to ~85% of all prominent syllables, which by far exceeded the highest maximum value produced by the native group (~60%).

Equally idiosyncratic was the distribution of the rising tones: their percentage was as low as ~3% in the realizations of the below-average male informant (vs. ~20% in the other groups). With respect to the other categories, the subject in question remained within the boundaries set by the general patterns.

Average group

Gender-dependent phenomena in the average group were not numerous across the two speech types.

In read speech, the main differences between male and female speakers were discovered in the two most frequent categories, where the male test persons exhibited a clear preference for L*H accents (~40%) over the H* accents (~32%). For female speakers the pattern was reverse: they employed the simple high targets more extensively (~40%) than the rising tones (~33%).

Spontaneous speech realizations exposed a similar, yet less prominent, tendency. The male subjects of average aptitude here made almost equal use of H* and L*H events (~36-38%,). Female speakers nearly replicated this distribution pattern, as far as the simple high targets are concerned (~40%). However, the percentage of rising tones in female speech was lower (~29%). Two other categories exposed cross-gender differences in spontaneously produced utterances, in a way that male speakers used more L* accents in their speech (~17%) than did the female speakers (~13%), but the number of H*L tokens was visibly lower in male productions (~5%) than realized by female subjects (~10%).

Low-aptitude group

Finally, low-aptitude subjects displayed gender differences in the realization of several categories as well. In particular, L*H accents were more typical of female-speaker utterances (~37%) than for male productions (~28%) in read speech. In the same speech variety, H*L events were, on the other hand, produced in greater numbers by the male subjects (~18%). The female part of the group applied it to just about a tenth of all accented syllables.

There were also three major mismatches in spontaneous speech. Firstly, the male speakers realized about a half of all investigated syllables with the H* accent, whereby this percentage went down by about 10% as spoken by the female test persons. The second difference was found in the L*H category, which was visibly more characteristic of the female subjects (~38%).

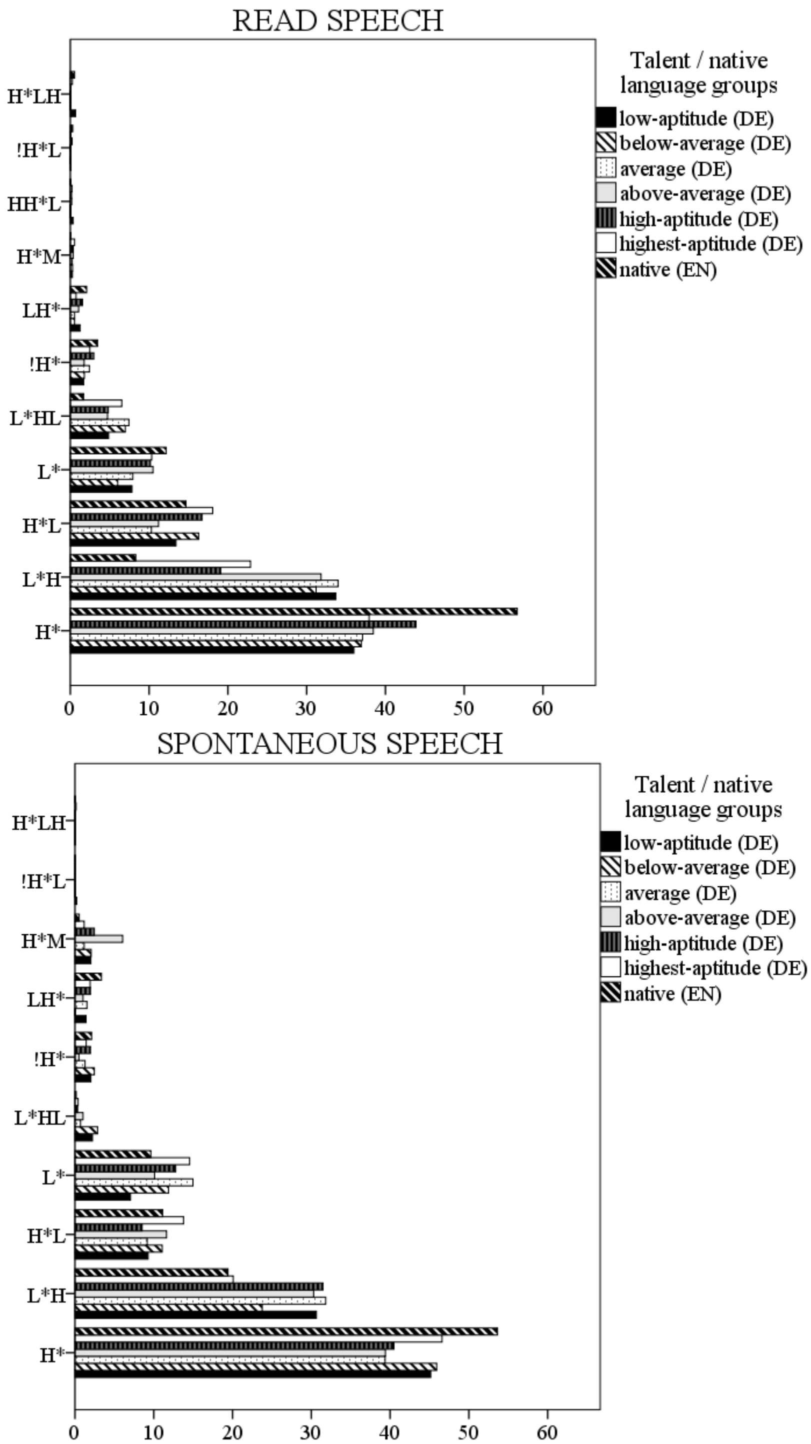


Figure 6.11: Distribution of pitch accents in all aptitude groups by speech type – read versus spontaneous realizations in the English-language part of the experimental corpus . On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed pitch accents

In male realizations this category was proper to about a fifth of prominent syllables in the current investigation. H*L events exposed the third major discrepancy across the genders: this category was observed in a larger portion of accented syllables in male speech (~12%), whereas it was only common for ~5% of test data in the female realizations.

Native subjects versus German groups

The main objective of the current investigation was to compare the performance of the native English control group with the L1 German subjects.

By applying the factors of speech type and gender simultaneously, we could see in more detail, which of these speakers were closest to the native subjects with regard to the distribution of pitch accents.

We could see therefore that both male and female native speakers applied simple high targets to a larger number of tokens in both speech types than most German groups. The only exception was the idiosyncratically high percentage of H* events in the realizations of the below-average male speaker, which, though breaking the pattern formally, was more of a sign of individual variability than a real steady trend. It should also be noted about the H* category that female native speakers did not depart as much in performance from the German subjects as did the male L1 English informants. The percentage of H* accents was ~10-20% greater for the latter and only ~2% - for the former, as opposed to the corresponding German speaker productions.

The general-corpus trend in L*H pitch accents was also scrutinized at this stage of analysis. As a result, we could consistently confirm the tendency of highest-aptitude and high-aptitude female speakers to suppress the employment of the rising tones when speaking English, which made their realizations closer to the ones produced by native speakers.

In fact, in spontaneous speech the highest-aptitude and high-aptitude female informants only marginally differed from the native female test persons (~20-24% of L*H accents in all of the three groups). Notably, the below-average female subjects demonstrated quite similar values (~26%). For the above-average and average subjects, the amount of L*H tokens constituted over a third of the prominent syllables; and low-aptitude test persons displayed an absolute maximum in this condition – about 38%.

Read speech realizations established a greater distance between the female native speakers and the L1 German female subjects in that the frequency of occurrence of the rising tones was ~5% lower in the native sample as opposed to the closest values rendered by L1 German speakers – the high-aptitude and highest-aptitude test persons (~17-20%). All the remaining groups demonstrated gradually increasing percentages of L*H accents with their increasing level of pronunciation aptitude: above-average and average (~25-30%), below-average and low-aptitude (~37%).

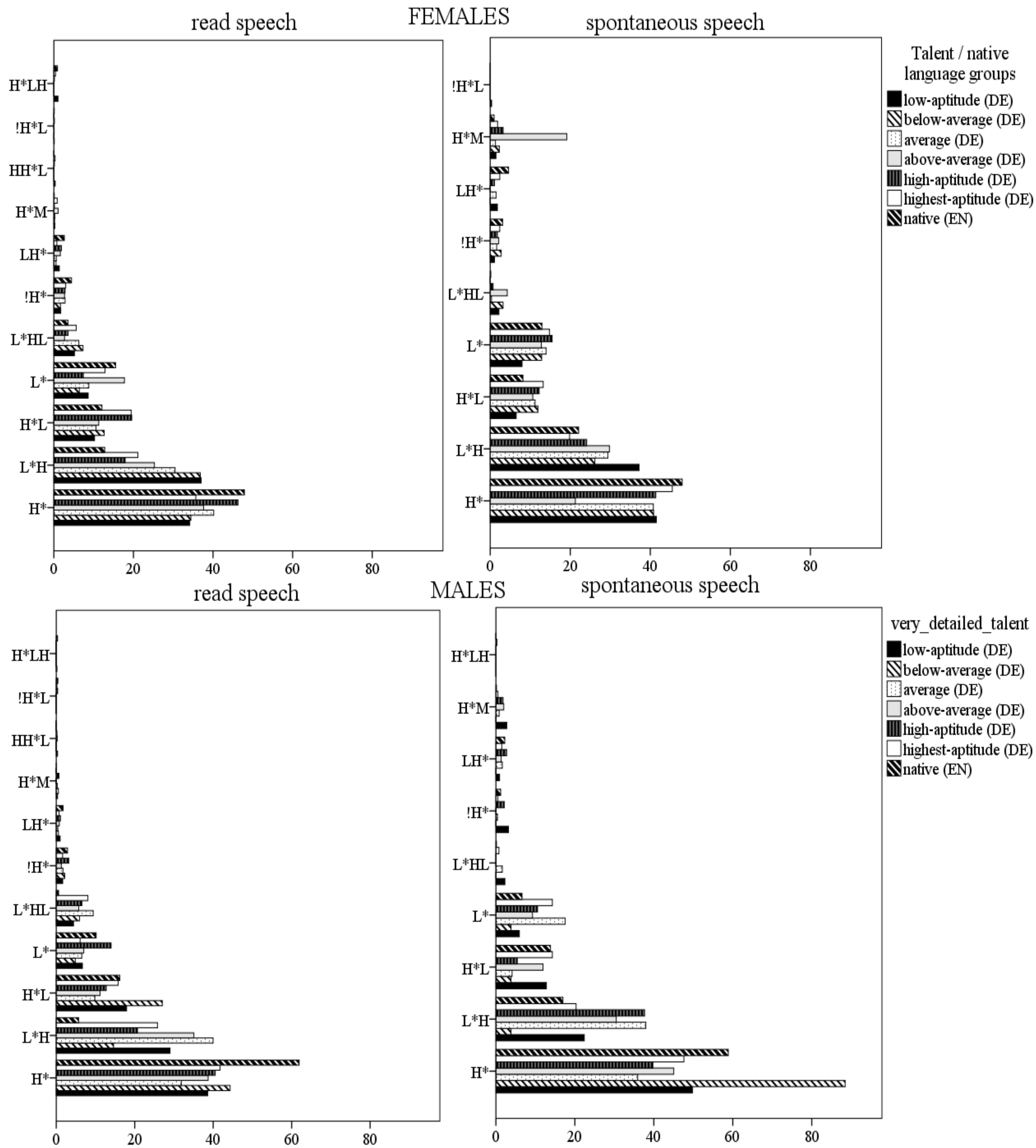


Figure 6.12: Distribution of pitch accents in all aptitude groups by gender – male versus female realizations – in read and spontaneous speech in the English language part of the experimental corpus. On the x-axes – frequencies of occurrence within a group in per cent; on the y-axes – the observed pitch accents

The distribution of rising tones in male realizations was similar to the female groups in that the native speakers in general produced smaller numbers of L*H targets than the L1 German subjects. This difference was, as with the female informants, more visible in read speech.

Here, the native male test persons only realized ~6% of all prominent syllables with a rising tone, whereby the smallest percentage registered for the L1 German speakers was ~16%, which was found for the below-average subject. His performance was followed by the high-aptitude, highest-aptitude and low-aptitude (~20-28%) groups. Finally, the proportion of L*H events was highest in the above-average (~36%) and average group (~40%) realizations. As seen from the above description, the dependence between the amount of rising tones and pronunciation aptitude discovered for female speakers was not confirmed for male read speech.

The same was also true for the spontaneously produced utterances. Moreover, the below-average speaker once again performed contrary to the overall trend by exhibiting a lower percentage of rising targets than the native speakers (~4% vs. ~17%, respectively). The remaining German groups had a greater number of L*H accents in their productions than the native speakers. Thus, the highest-aptitude and low-aptitude informants employed the rising tokens slightly more extensively (on ~20-23% of all prominent syllables). The other groups had visibly higher values in this respect (~32-38%).

H*L was one of the categories where native speakers did not deviate from the German subjects in performance.

Female productions indeed presented quite clear-cut patterns in this sense. In read speech, most groups, including the native informants, had ~12% of all syllables carrying a falling target. The exception was presented by the highest-aptitude and high-aptitude groups with larger percentages (~20%). Most spontaneously produced utterances also included ~12% of H*L accents in them. Only native and low-aptitude female test persons did not comply with this trend by producing fewer H*L tokens (~7-9%).

Male speech did not reveal any clear trends in the distribution of the H*L pitch accents. The main observations concerned the aberrant behaviour of the below-average subject in read speech, who produced ~27% of all prominent syllables with a falling tone. The percentages in all the remaining groups were lower (~13-19%). In spontaneous speech, most groups had approximately equal proportions of falling tones (~16%), but the high-aptitude, average and below-average groups displayed lower values (~4%).

More regularity in female speech was also found for the simple low targets. In read utterances, native, highest-aptitude and above-average informants employed the L* category slightly more extensively than the other groups (~16-19% vs. ~9%). Spontaneous speech only singled out low-aptitude female subjects with their lower percentage of low targets than average (~10% vs. ~15%).

Male productions once again did not yield any steady trends. In read speech, only the high-aptitude and native groups demonstrated higher percentages of L* targets (~13-16% vs. ~7% in the other groups). In spontaneous realizations, ~12-18% of all ip's were realized on a low pitch. However, native, below-average and low-aptitude speakers demonstrated lower percentage values (~4-6%).

As mentioned previously, L*HL events formed a frequent category in read speech only, and these productions did not show any patterns of native speaker performance as contrasted with the L1 German groups genderwise.

Figure 6.12 presents graphical and statistical evidence for the empirical observations regarding the distribution of pitch accents in read and spontaneous speech by gender.

We have hereby analyzed the distribution patterns of all the most frequent categories in the English part of the experimental corpus and compared the German groups with native speakers under all of the foreseen conditions. We will further proceed with the analysis of boundary tones in the same general fashion.

6.1.2.2 Boundary tones

At the final stage of the phonological investigation of ToBI categories, we analyzed the distribution of boundary tones in the English part of the experimental data. This effort resulted in some pertinent observations on the performance of both native and non-native speakers of varying pronunciation aptitude.

General corpus analysis

In general, we discovered the same categories as being frequent in the corpus, as the ones registered for the German data. Thus, the largest number of tokens pertained to the full low boundary L-L% category (~40-50%), the H-H% events (~27-35%) and the typically German default boundary % units (~15-20%).

Low full boundary tones L-L%

The first one of these categories, i.e. L-L%, was clearly more typical of the native speakers, who realized about a half of intonation phrases at a low pitch level. Highest-aptitude, high-aptitude and, incidentally, below-average subjects followed with ~39-42% of such events. Finally, L-L% type was least common for low-aptitude, above-average and average speakers (~35%).

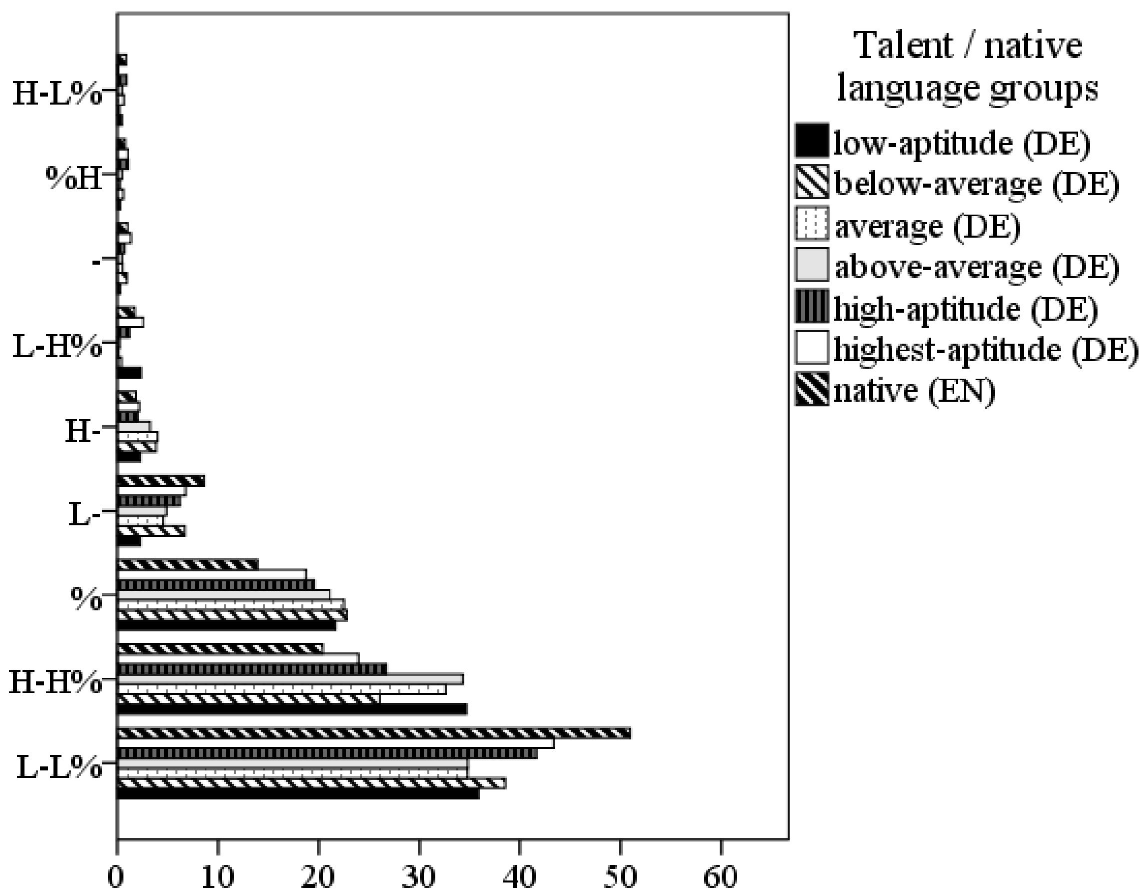


Figure 6.13: Distribution of boundary tones in all aptitude groups in the English language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

High full boundary tones H-H%

The full high boundaries H-H% presented an opposite trend in the data. Here, native speakers produced the smallest number of intonation phrases on a high pitch level (~20%). Highest-aptitude, high-aptitude and below-average subjects were once again closest to the native group in this respect, by increasing the percentage of H-H% events by only ~3-5%. The remaining groups clearly differed in performance in this category, as they realized ~35% of all intonation phrases with a full high-level boundary.

Default boundaries %

A peculiar fact in general was the presence of default boundary tokens in the native speech. As discussed in Chapter 5, this boundary type is not part of the original ToBI classification for English. However, we labelled the data based on the combined inventory for both languages, and ~15% of all boundary events in the native-speaker realizations fully complied with the definition of the default boundary category as the IP-final unchanging pitch. For all the German groups this category was, quite logically, more frequent. All the German groups exhibited similarity in their percentage values (~19-22%).

Intermediate low L- and high H- boundary tones

Fairly frequent was also the low intermediate boundary L-. Here, the pattern would be similar to the previous categories, where the distance in percentages rises with the decreasing aptitude level, except for the general idiosyncratic behaviour of the below-average group, who exhibited values almost equal to those of the highest-aptitude subjects. Otherwise, the percentage of tokens in this category gradually decreased from the native speakers (~9%) to the low-aptitude group, who displayed a visible drop in numbers (~2%).

The intermediate high boundary H- accounted for ~2-4% of all ip's in all groups, with a slight preference for this type by the above-average, average and below-average speakers[16].

As the next step we applied the gender factor to the data with the hope that it would provide a more detailed insight into the distribution of ToBI boundary tones.

16 Figure 6.13 is to be referred to for an illustration of boundary tone distribution in the whole experimental corpus.

Gender-specific analysis

Low full boundary tones L-L%

During the gender-specific analysis, it became apparent that male native speakers were mainly responsible for the significant difference between native and L2 subjects in the realization full low boundaries L-L%. Over a half of the intonation phrases carried an L-L% marker in the male native group productions, whereby in the L1 German male groups the percentages did not go over ~42%, demonstrated by the highest-aptitude and low-aptitude test persons. The proportion of low full boundaries in the remaining groups was ~35%.

For the female speakers the general number of L-L% events did not go over ~45%, demonstrated by the female native speakers, as well high-aptitude and highest-aptitude female groups.

Female speakers also produced a more “regular” pattern with regard to the aptitude factor. Apart from the three-group cluster, which we have just mentioned, above-average and average informants also demonstrated similar performance by producing ~35-37% of all intonation phrases on a low final pitch level. Finally, the low-aptitude speakers only used the L-L% events in ~31% of all cases under analysis.

High full boundary tones H-H%

In the full high boundary category H-H% female subjects displayed similar clustering as was the case with the L-L% events. More specifically, the low-aptitude speakers stood out from the other groups with the highest percentage in this category (~38%); the above-average, average and below-average groups followed with ~29-32%; and the highest-aptitude and high-aptitude, as well as the native subjects formed the final cluster with about a fourth of all intonation phrases realized on a high pitch level.

Less homogeneous was the picture in the male groups, as far as H-H% events are concerned. The below-average male speaker performed here in a similar way as the native group due to the fact that he produced almost an equal number of intonation phrases with an H-H% token - ~15% vs. ~17% in the native realizations. In the remaining groups the proportion constituted from ~25% to ~35% of all intonation phrases.

Default boundary tones %

The default % boundaries were produced by the female speakers in a fairly regular way. Although the percentage of tokens in this category was lowest in the native group (~16%), it was very close to the highest-aptitude and high-aptitude (~18%) speakers. The values in the remaining groups were only slightly higher (~21-23%).

Male speakers produced a similar pattern, yet the difference between the native subjects and the other groups was greater - ~10% against ~19-26%, respectively. The maximum percentage value pertained to the below-average speaker in this respect.

Low intermediate boundary tones L-

As far as the low intermediate L- boundaries are concerned, most female groups produced an almost equal number of intermediate phrases on a low pitch level, which accounted for ~6-7% of the test data. The exception was the low-aptitude group, for whom this category was far less frequent (~2%).

In male realizations the below-average speaker once again performed in a similar fashion as the native speakers, by producing about a tenth of all intermediate phrases with an L- target. Highest-aptitude and high-aptitude speakers were also close to the native group in this respect, with a minor two-percent drop in percentages. In the above-average, average and low-aptitude groups this proportion was reduced by about half (~5%).

High intermediate boundary tones H-

The distribution of intermediate high boundaries did not yield any clear trends, with percentages of H- tokens of ~2-5% in all groups.

Figure 6.14 is a graphical representation of boundary tone distribution in the English part of the experimental corpus by gender.

The application of the second factor, i.e. the speech type, allowed us to look at the data from still another angle.

Corpus analysis by speech type

Low full boundary tones L-L%

In read speech, the distribution / frequency of occurrence of the full low boundaries presented a rather clear pattern. More specifically, the native speakers were singled out here as the group with the highest percentage of L-L% tokens (~62%). The highest-aptitude and high-aptitude informants were closest to the native speakers in this respect by realizing the greatest number of intonation phrases of all L1 German groups at a low pitch level (~50%). All the remaining groups produced smaller amounts of L-L% events (~39-45%).

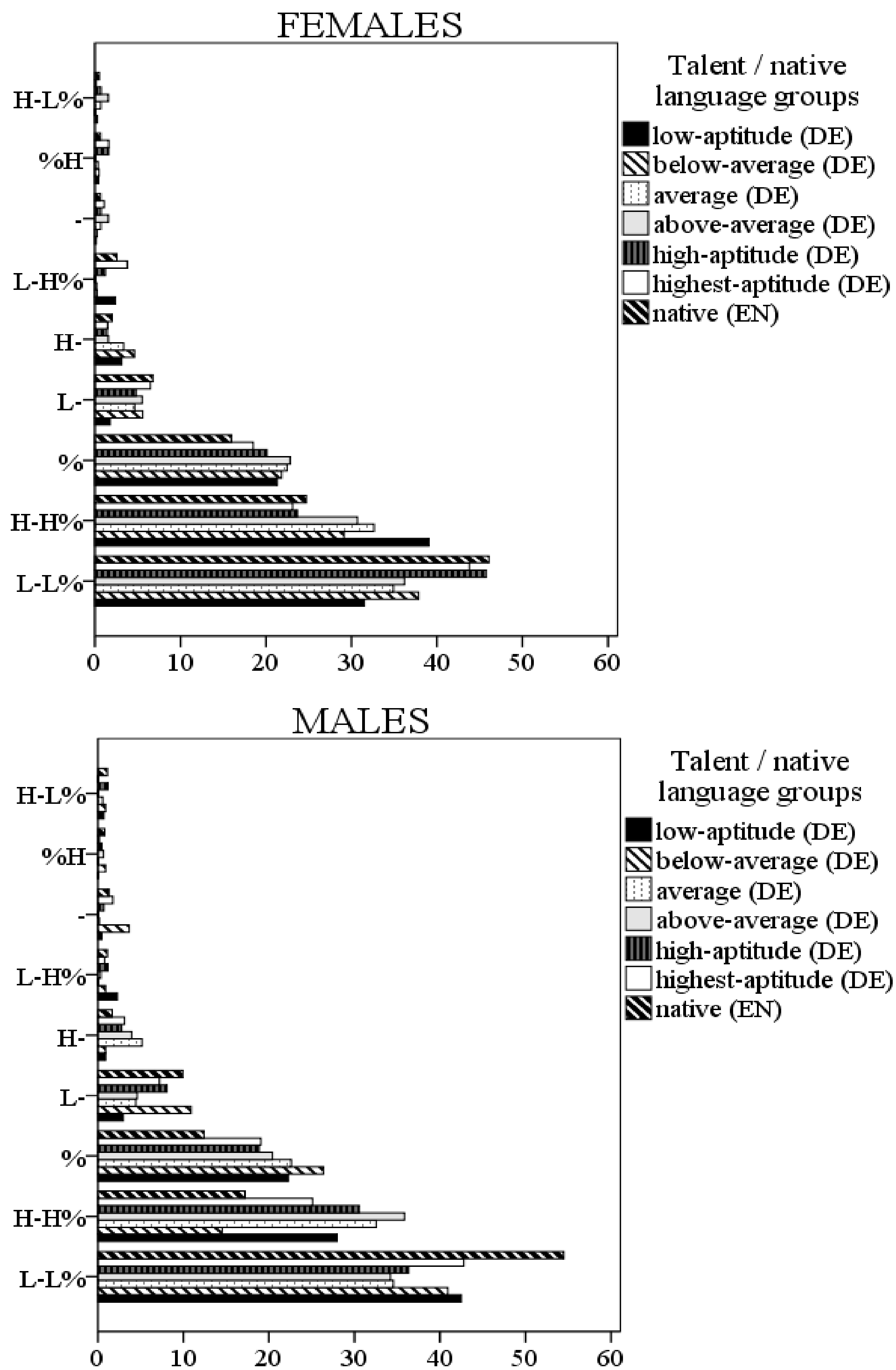


Figure 6.14: Distribution of boundary tones in all aptitude groups by gender – female versus male realizations in the English language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

Quite a different picture was discovered for L-L% boundary tones for spontaneous speech. Native speakers still, as in read speech, had the greatest number of low-pitch intonation phrases (~35%).

In this condition, however, such performance was most closely approached by speakers of highest and low pronunciation aptitude (~30%). The remaining subjects exhibited almost equal numbers of L-L% targets in their spontaneous speech (~20-22%).

Low full boundary tones H-H%

For the full high boundaries in read speech, we could discover a similar, though reverse, aptitude-dependent pattern observed for the L-L% category. As was the case with the full low boundaries, the native speakers could be separated from all German groups due to the fact that they produced fewer IP's at high pitch (~15%). The highest-aptitude and high-aptitude speakers were once again in closest proximity to the native group in that they produced the smallest number of H-H% tokens of all German subjects (~20%). In their turn, the below-average informants approached the two higher aptitude groups (~25%). The remaining groups realized approximately equal numbers of intonation phrases with an H-H% event (~30-32%).

Spontaneous speech unveiled a two-cluster pattern in the distribution of the H-H% category. In the first cluster, the native, highest-aptitude and below-average groups exhibited similar percentages of high full boundaries (~28%). The remaining informants employed H-H% tokens more extensively – on ~36-38% of all IP's.

Default boundary tones %

With respect to the default boundaries, most German groups performed similarly in read speech, producing % targets with almost equal frequency of occurrence (~15%). Only the low-aptitude group departed from this trend with about a fifth of all intonation phrases ending on an unchanging pitch. Lastly, the native speakers differed from all German subjects in this category, as in their speech only ~11% of all IP's carried a default boundary tone.

Native speakers also had the smallest number of % tokens in spontaneous speech (~19%), as compared to most L1 German groups (~25-30%). The average and below-average subjects were the two groups applying default boundaries most frequently (~36-37%).

Intermediate low L- and high H- boundary tones

The low intermediate boundaries were quite evenly distributed in both speech types, applied to up to a tenth of intermediate phrases. Only several groups marginally deviated from this overall trend: in spontaneous speech – low-aptitude and average informants (with ~5%), as well as native speakers (~12%); in read speech – low-aptitude and above-average test persons (~with 3-5%).

The final frequent boundary tone category, the high intermediate boundaries, did not reveal any clear inter-group peculiarities. They were characteristic of ~2-5% experimental cases.

Distribution of boundary tones in the English-language part of the experimental corpus is illustrated in Figure 6.15.

The analysis of ToBI boundary tone distribution in English was finalized by the simultaneous application of both the speech type and the gender factor to the data.

Corpus analysis by speech type and gender

Highest-aptitude group

The first striking observation in the analysis of boundary tones by speech type and gender was the degree of agreement in the highest-aptitude group: no significant differences were found between the male and female speakers in this respect, except for a slight 2-percent mismatch in the low and high full boundary distribution in spontaneous speech.

High-aptitude group

High-aptitude subjects did exhibit a few differences genderwise. Thus, the percentage of full low boundaries L-L% was lower for the male speakers in both speech types. In read realizations there was a drop from ~55% to ~48% and in spontaneously produced utterances – from ~27% to ~19%, as compared to female speakers. In spontaneous speech male speakers also realized a larger number of intonation phrases with a full high boundary H-H% (~42% vs. ~34% in the female group). Finally, male subjects employed the intermediate high boundary (~7%) in spontaneous speech more extensively than the female speakers, who produced a minor number of ip's at high pitch level.

Above-average group

Several gender-specific peculiarities were found when we analyzed the distribution of boundary tones in the above-average group. Thus, one of the biggest differences concerned the frequency of occurrence of the H-H% category in read speech. The was about a 6-percent distributional mismatch in this condition, whereby these tokens were more frequent in male speech as opposed to the female realizations (~35% vs. ~29 %).

There were also some visible discrepancies in spontaneous speech with respect to the full low boundaries and the default boundary events. The latter category was more common for the female informants (~40% vs. ~27%); the male test persons, on the other hand, employed the L-L% events more extensively (~24% vs. ~12%).

Average group

In the average group the only major difference was observed in spontaneous speech, where the female subjects produced ~5% more intonation phrases with a high full boundary than the male speakers (~40% vs. ~35%).

Below-average group

Inter-gender differences in the below-average group were discovered for most conditions, especially in spontaneous speech.

More specifically, there were only ~6% of H-H% and L-L% tokens in spontaneous productions of the male subject. The corresponding percentages produced by the female group were visibly higher (~30% and ~25% for the two above categories, respectively).

The number of default boundary tones as realized by the male subject in spontaneous speech was abnormally high (~70%), which was reduced by half in the intonation phrases rendered by the female speakers (~32%).

Intermediate low boundary was the final category where we found cross-gender differences in spontaneous speech. Thus, such tokens were almost absent from the male subject's speech; the female informants, on the contrary, applied this category on about a tenth of all intermediate phrases.

The major differences in read speech involved the just mentioned intermediate low boundaries and the H-H% category. Opposed to the observations made for spontaneous speech in the corresponding category, the male subject had a greater number of low intermediate boundary targets than did the female speakers (~12% vs. ~3%). Finally, the percentage of H-H% tokens was clearly lower as produced by the male subject than the female counterpart group (~15% vs. ~28%).

Low-aptitude group

There were quite a few mismatches in the distribution of the ToBI boundary tones in the low-aptitude group across the two genders. In read speech male speakers realized a visibly larger number of intonation phrases with an L-L% token (~45% vs. ~34%). Female subjects, on the other hand, employed the high full boundary more extensively in this speech type than the male part of the group (~35% vs. ~23%). A similar pattern was found in spontaneous speech. Here, there was a significant drop in percentages of L-L% tokens between the genders: ~37% in male realizations and ~25% in female speech. In the high full boundary category the difference was also significant: ~46% of intonation phrases were spoken with a high final pitch level by the female test persons, whereby with the male speakers this was only the case with a third of all IP's.

Native subjects versus German groups

The application of the gender and the speech type factors simultaneously allowed us to carry out a detailed examination of the trends discovered for the entire corpus, as well gender-specific and speech type tendencies, in native subjects' performance as opposed to all of the L1 German groups.

Low full boundary tones L-L%

We could therefore confirm the tendency of the native speakers to employ the low full boundaries more extensively than the German informants in both speech types. Moreover, we could also confirm the fact that the difference between male native speakers and the L1 German male test persons was greater than the corresponding contrast in female speech.

In male productions in read speech, there was about a 15-percent mismatch in the frequency of occurrence of the L-L% category between the native subjects (~65%), on the one hand, and most German speakers (~47-50%), on the other. The above-average and average groups had the lowest number of low full boundary tokens in their speech (~40%).

Spontaneous speech unveiled unexpected behaviour by the low-aptitude male subjects, who applied the L-L% events to an almost equal number of intonation phrases as the native speakers (~38%). The remaining groups had fewer instances of the L-L% category in their speech.

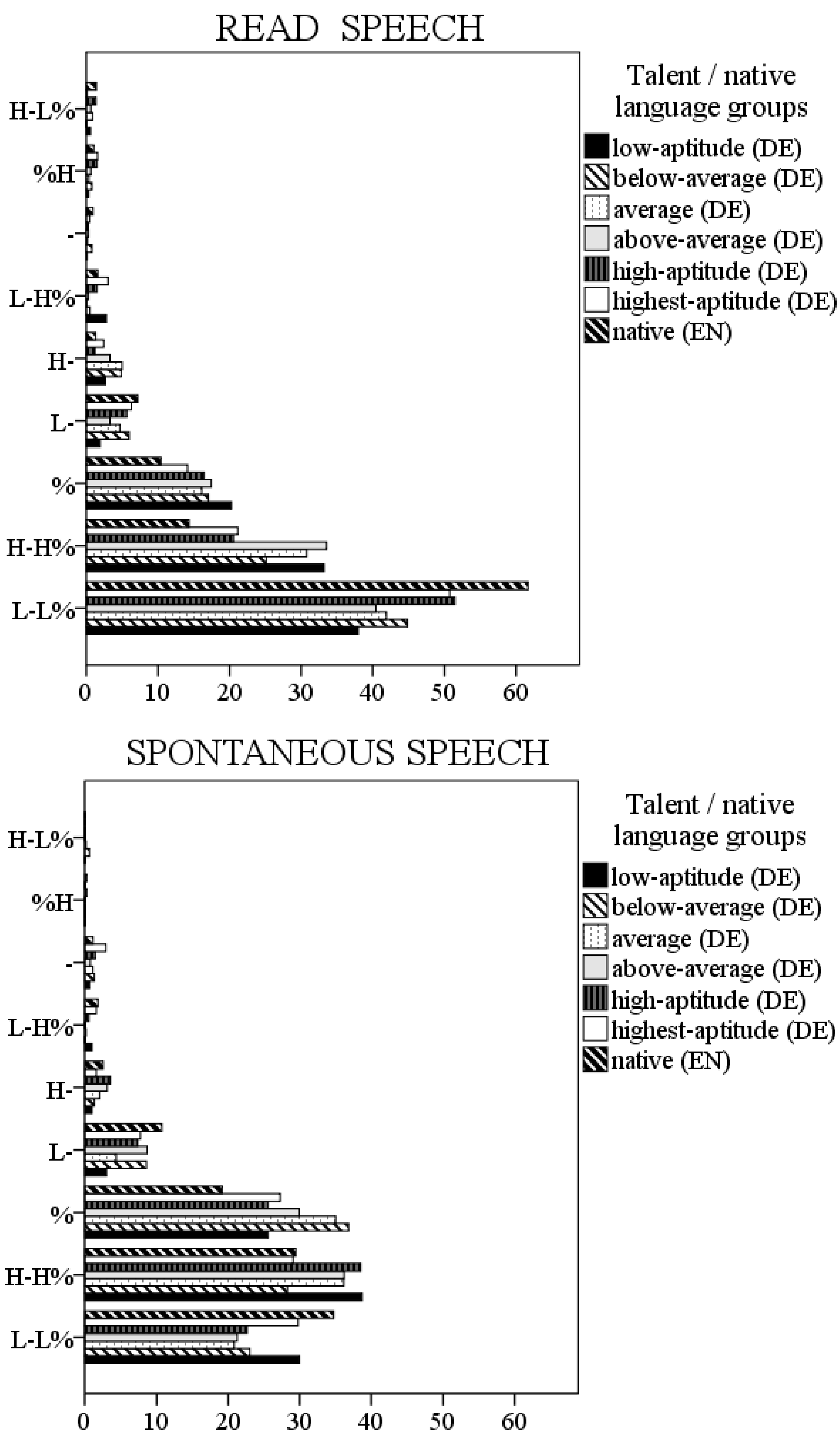


Figure 6.15: Distribution of boundary tones in all aptitude groups by speech type – read versus spontaneous realizations in the English language part of the experimental corpus. On the x-axis – frequencies of occurrence within a group in per cent; on the y-axis – the observed boundary tones

The percentages decreased from highest-aptitude (~31%) to above-average, high-aptitude and average (~20-24%) and, finally, below-average (~6%) speakers.

Female productions once again allowed of an aptitude-dependent interpretation. In read speech, native group performance was closely approached by the highest-aptitude and high-aptitude groups (58% vs. ~50-54%, respectively). In most of the other female groups the percentage of low full boundary targets lay by ~44%, the only exception being the low-aptitude informants (with ~33%).

Looking at spontaneously produced utterances, we could state the fact that most female groups were really close to the native speakers, as far as the frequency of occurrence of the L-L% events is concerned (~25-28% for the German subjects; ~30% for the native group). The two exceptions to this observation included the average and above-average informants, applying the low full boundary type to a smaller number of IP's in spontaneous speech (~21% and ~12%, correspondingly).

High full boundary tones H-H%

We could also corroborate some of our earlier observations regarding the full high boundary category, when applying both the speech type and gender factors to the data.

Overall, female speech could be described as more homogeneous and regular in this category than male productions. In read utterances, for instance, almost no distinction could be drawn between the native, highest-aptitude and high-aptitude speakers' realizations with respect to the frequency of occurrence of the H-H% events (~20%). Most other groups applied the full high boundary to about a third of intonation phrases in read utterances. This percentage was, however, higher within the low-aptitude informants (~37%).

Even more regularity could be found in spontaneous speech. In this condition, most groups had ~30-33% of all boundary targets pertaining to the H-H% category. Slightly deviating from this pattern, the highest-aptitude group employed the events in questions less frequently (~27%). Finally, the low-aptitude test persons appeared to be clear outliers by producing a visibly larger number of intonation phrases at high pitch level (~46%).

Turning our attention to male speech, we could not notice any clear-cut tendencies. Nevertheless in read speech the native speakers could be separated from the German groups due to their smaller numbers of H-H% tokens (~10%). Unexpectedly, the below-average subject exhibited closest values to the native group in this condition (~15%), which was followed by highest-aptitude and high-aptitude subjects producing still higher percentages of H-H% tones (~20%). Finally, the low-aptitude and above-average test persons employed the full high boundary category most extensively – with ~30-35% of IP's ending at high pitch in read speech.

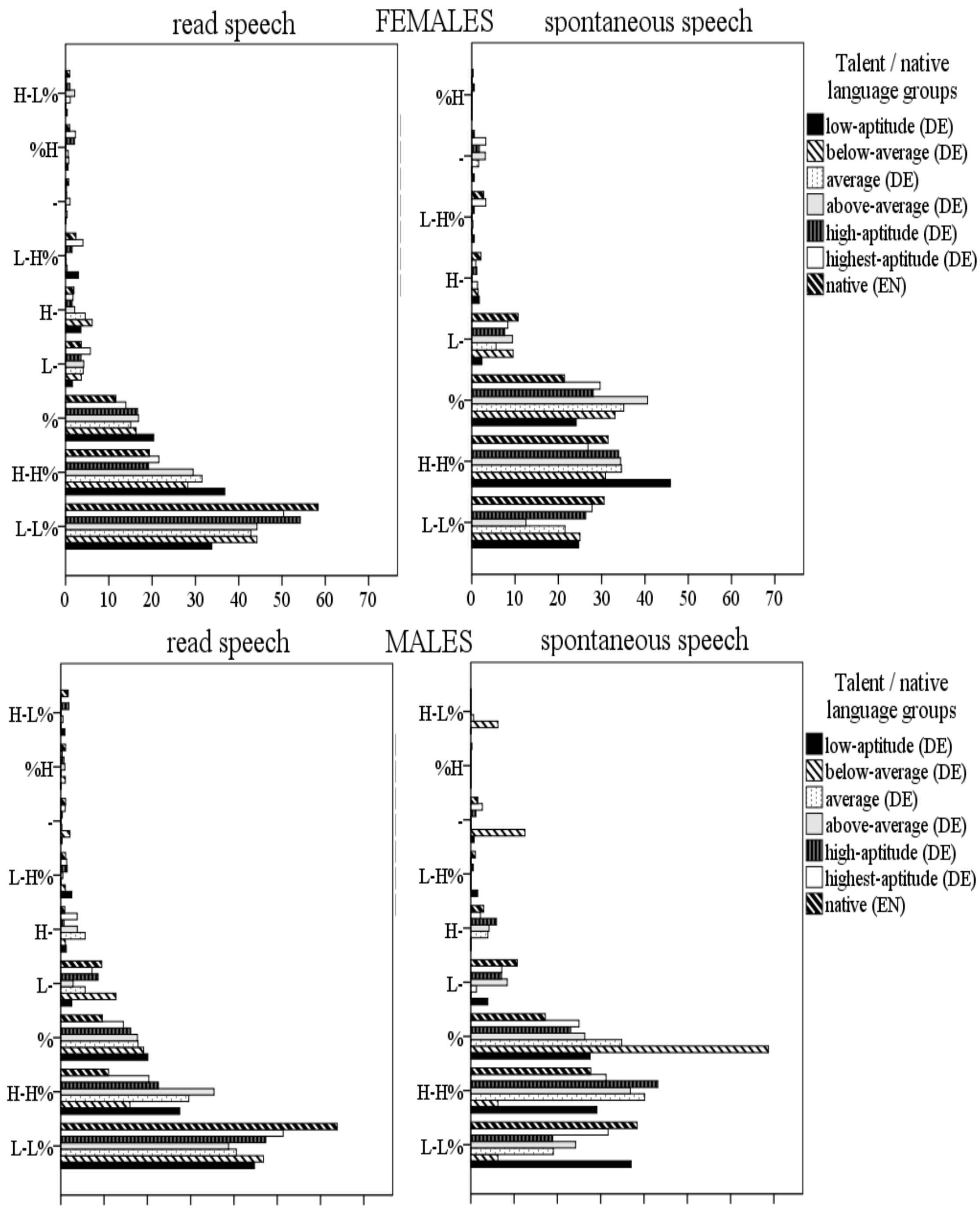


Figure 6.16: Distribution of boundary tones in all aptitude groups by gender – male versus female realizations – in read and spontaneous speech in the English language part of the experimental corpus. On the x-axes – frequencies of occurrence within a group in per cent; on the y-axes – the observed boundary tones

Spontaneously produced utterances also contained some clear cross-group mismatches in male productions.

Thus, the below-average subject displayed a conspicuously low percentage of H-H% targets in his speech (~6%), whereas the native, highest-aptitude and low-aptitude informants applied the full high boundary in about a third of all cases in this condition. The number of H-H% events was highest in the remaining groups (~40%).

Default boundary tones %

The default boundary was the only category where male and female subjects displayed almost equal homogeneity in read speech. For male speakers, this was manifested by a lower general percentage of % events produced by native speakers (~10%) and a gradual increase in % tokens with the decreasing degree of pronunciation aptitude – from ~15% to ~20%. In female speech, the distance between native and German speakers was only marginal, and there was a similar pattern of slightly increasing percentages of % boundary tones from the native to the below-average subjects (from ~12% to ~17%, respectively). The low-aptitude informants deviated slightly with about a fifth of all IP's carrying a % target.

Spontaneous speech presented more cross-group mismatches in the given category.

First we examined female realizations in this regard. The lowest percentage of % targets found thereby in the native speakers' productions (~20%) was followed by a slightly higher amount produced by the low-aptitude subjects (~23%). The application of % events by most other groups was more extensive, characteristic of about a third of intonation phrases. Finally, the above-average female group had the greatest number of % tokens in spontaneous speech (~40%).

The most striking observation for male realizations was, unsurprisingly, related to the below-average subject's performance, who, as mentioned earlier in the section, applied the default boundary to ~70% of all intonation phrases in spontaneous speech. The percentages in most other groups lay at 25%, except for the average speakers with a greater proportion of % targets (~35%) and native male speakers, who employed the category in question most sparingly (~18%).

Intermediate low boundary tones L-

The low intermediate boundary category enjoyed a considerably regular distribution in female productions in both speech types. In read speech, almost all groups applied the given targets to an almost equal number of ip's (~3%), with the only exception of the low-aptitude speakers with a lower percentage (~1%). This latter group also differed from the other groups in performance in spontaneous speech by producing a similarly low number of L- tokens (~2%), whereas the other female subjects applied the category at issue more extensively – to ~7-10% of all full boundaries.

Contrary to the regularity of the female realizations, male productions presented a heterogeneous picture in the distribution of intermediate low boundaries. In read speech, we could register similar performance by the native, highest-aptitude, high-aptitude and below-average groups, who applied L- targets to an almost equal number of ip's (~9-12%). The remaining groups could be identified by the lowest proportion of L- tokens (~3-5%).

In spontaneous speech, native, highest-aptitude, high-aptitude and above-average speakers employed the L- category to similar percentages of intermediate phrases (~10%), whereby the remaining groups demonstrated lower values (~1-4%).

Intermediate high boundary tones H-

The number of H- events in all groups and in both speech types did not exceed ~4% of all the examined intermediate phrases. No clear tendencies or cross-group peculiarities could be registered in this regard.

Figure 6.16 reflects the gender-specific findings in the distribution of boundary tones in read and spontaneous speech.

In the current section, we presented the tendencies in the distribution of pitch accents and boundary tones, both in the German and English parts of the experimental corpus. These results can be used at the subsequent stages of the current study, i.e. for examining the phonetic properties of each of the stipulated ToBI categories.

From a phonological perspective, however, we undertook an additional investigation of these phonological categories on a so-called meta-level. Bearing in mind that part of the data comes from read speech, namely, the utterances pertaining to two texts, "The North Wind and the Sun" and the abstract from M. Twain's "Mrs McWilliams and the Lightning", we wished to analyze the intonational interpretation of these texts, based on the corresponding ToBI transcriptions. Comparing these transcriptions would therefore allow us to see how uniform / variable the text interpretation was in each aptitude group. That, in turn, would possibly lend some new evidence of a possible correlation between intonational variability and pronunciation aptitude.

The outcome of such a phonological investigation will be presented in Section 6.2.

6.2 Tendencies in general prosodic text interpretation

The second stage of the phonological analysis consisted in exploring the distribution of ToBI events on the text level in order to see how uniform, or variable, the prosodic interpretation of the text by the native and the German speakers was. Two texts were analyzed in this regard – an abstract from “Mrs McWilliams” by Mark Twain and “The North Wind and the Sun” fable. We first analyzed the whole corpora pertaining to each of these texts and, further on, applied the gender factor to the data, thus investigating the female and male realizations separately.

The Levenshtein distance calculation procedure

The statistical measure employed for the above analysis of prosodic text interpretation is commonly referred to as the Levenshtein distance (Levenshtein 1965). This procedure was thereby applied to compare all ToBI transcriptions of the two investigated texts.

Levenshtein distance measures the minimum number of corrections needed to transform one string into another. Therefore, any insertion, deletion, or substitution of a single character increases the distance between each pair of strings at issue. For example, “bit” and “kits” have a Levenshtein distance of 2: substitution of “b” by “k” and the insertion of “s” at the end of the second word. The given statistical method has been commonly used for approximate string matching, i.e. finding similar character sequences. We therefore assumed that the Levenshtein procedure would be suitable for comparing prosodic transcriptions in the current investigation in that it would allow us to see the cross-group differences / similarities in the global text interpretation.

In practical terms, we analyzed equal-length syllable sequences corresponding with the two texts. Each of the syllables, in turn, would or would not carry a ToBI event – a boundary tone or a pitch accent. It is the strings of ToBI tokens (substituted by a single-character code for convenience) intermittent with unstressed syllables (coded with a specific character as well) that we processed. Should a syllable carry both a pitch accent and a boundary tone by any one or more speakers, the corresponding syllable was printed twice in the transcription. Thus, we compared prosodic transcriptions produced by all the subjects in each particular aptitude group. The cross-speaker differences in Levenshtein distance terms would then only result from the substitution of characters, i.e. two ToBI events or an unstressed syllable with a ToBI instance. Due to the insufficient number of the calculated Levenshtein distance indications in the above-average and the below-average groups when split by gender, only one string per group, which made these speakers incomparable to the other groups statistically, we excluded the gender-specific above-average and below-average samples from the analysis.

The North Wind and the Sun

To facilitate the analysis that will follow and as a general remark, we ought to reiterate that the prosodic transcriptions we analyzed were of equal length, and for German they were constituted of strings of 193 characters, whereby the English version of the text resulted in slightly longer sequences of 203 symbols.

Relative to these general numbers, the number of cross-speaker differences in each of the groups in English and German fluctuated between ~20 and ~100, i.e. between a fifth and a half of the possible discrepancies.

There were, however, some quite distinct cross-language peculiarities in the degrees of uniformity / heterogeneity in text interpretation. Thus, all German samples were much more concentrated around the mean values and contained fewer noticeable cross-group contrasts: the majority of tokens were found between Levenshtein distance values of ~40 and ~60. Only the highest-aptitude informants demonstrated quite a number of instances of greater cross-speaker differences, lying beyond the virtual average upper limit of ~60, but not higher than ~80. The low-aptitude group realizations also yielded some higher than average Levenshtein values (up to ~70).

In the English part of the data the distribution of Levenshtein distance values unveiled a more heterogeneous picture. First of all, the tokens were more scattered (between ~20 and ~100 on the Levenshtein scale). Secondly, there was also somewhat of a layering of groups in this condition. The average and native groups demonstrated values towards the relative minimum (between ~60 and ~20); the low-aptitude speakers, on the other hand, exhibited greater inter-speaker disagreement in text interpretation (between ~70 and ~100); finally, the remaining test persons took an intermediate stance by producing tokens in the range of ~50-75.

The ANOVAs carried out to test the above empirical observations attested to the significance of cross-group differences in both the English and German parts of the corpus (p=.000)[17]. Unsurprisingly, most of the Scheffé post-hoc comparisons also produced statistically significant results. As expected, in German the highest-aptitude and low-aptitude groups with their higher Levenshtein values were singled out as deviating statistically from the all the other samples (p<.20), but not the above-average realizations (p>.31). The above-average speakers, in their turn, also rendered a statistically different sample as opposed to the other groups (p<.01), excluding the just mentioned highest-aptitude and low-aptitude informants (p>.31).

The only steady trend in the English-language Scheffé tests concerned the low-aptitude speakers who significantly differed from all the other subjects as to their Levenshtein distance values (p=.000), with the only exception of the highest-aptitude group (p=.052). This fact was most probably accounted for by the visibly higher value range in the low-aptitude group than the rest

[17] p<.05 was taken as the significance threshold.

of the subject pool. The highest-aptitude subjects tended towards higher than average level, but their outcome was only distinguishable from that rendered by the average, below-average and native groups ($p<.02$). The last three significant contrasts in this experimental condition included the ones between the high-aptitude informants versus the average and native groups ($p=.000$) and the difference between the low-aptitude and native speakers ($p=.000$).

Figure 6.17 renders the graphical representation of Leveshtein distance values in "The North Wind and the Sun" corpus, whereby Appendix Tables 29, 30 illustrate the results of the statistical Scheffé tests in this relation.

By splitting the data gender-specifically, we obtained a more detailed insight into the data. However, the general tendencies discovered for the whole corpus appeared to persist in male and female productions even taken separately. Moreover, the cross-group differences registered for the whole corpus became clearer. The only exceptional condition in that respect was German male speech: only the average-group productions yielded a trend towards lower Levenshtein values (~40-50), whereby the other groups shared the same value space (~50-65). The female productions in German were also, for the most part, to be found within the same Levenshtein distance range (~40-60), although the average and high-aptitude informants demonstrated the majority of tokens in the lower part of this range (~40-50) and the low-aptitude group approached the relative maximum (~50-60). Conspicuously departing from the performance of the other female groups, the highest-aptitude subjects demonstrated a number of tokens on a higher than average level (~60-80) in their L1 read speech.

Taking due account of all the cross-group peculiarities in their L1 speech, it should be pointed out that more variability and mismatches were found in the English-language realizations. In both male and female data, pertaining to the average and native groups, we found generally more agreement in prosodic text interpretation, hence lower Levenshtein values (~30-55 for females; ~20-50 for males), than that expressed by the remaining speakers. For the latter, we discovered some gender-specific peculiarities.

Thus, the two higher-aptitude female groups appeared to show similar behaviour by producing higher values on the Levenshtein scale (~55-70) than those just mentioned for the average and native subjects. The greatest degree of cross-speaker disagreement expressed through the corresponding Levenshtein values was detected within the low-aptitude informants (~75-100); only several of their tokens were located lower on the scale (~40-55).

On the whole, comparable performance characterized the counterpart male groups: the high- and highest-aptitude groups produced intermediate-level values (~60) and the low-aptitude informants – the highest-lying tokens (~55-80). However, these values were more spread out than the respective female realizations. The values in the other groups lay between these extreme values in both female and male samples (~40-60).

As customary in the current study, we ran a series of statistical tests in order to test the above observations. It was stated thereby as a result of one-way ANOVAs that the cross-group differences in all language / gender conditions were significant (p=.000). The more specific Scheffé post-hoc tests rendered some expected as well as unexpected outcome. More specifically, the overall greater degree of disagreement in prosodic text interpretation by the highest-aptitude female subjects in German made this sample stand out statistically from the other ones (p<.01). Quite surprisingly, low-aptitude group realizations were also found to be deviating significantly from those produced by the other subjects (p<.03). In male German speech, only the average group with their lower Levenshtein values presented a statistical outlier as opposed to the remaining speaker pool (p<.03).

For the English data, part of the empirical analysis was also found to be valid. In the case with the female subjects this was mostly related to the performance of the average group, who consistently deviated from the other samples (p<.01), apparently due to the lower Levenshtein values produced by them. The native-speaker productions, also characterized by a relatively high level of cross-speaker agreement in intonational text interpretation, were only different from the average and low-aptitude groups in this respect (p=.000). This fact can be possibly explained by the presence of several tokens in a higher Levenshtein-scale range. The only other statistically significant contrast involved the high-aptitude and low-aptitude groups (p=.045).

In the utterances pertaining to the male subjects, the difference between the native speakers as opposed to the other groups was quite evident in the statistical sense (p<.01). Apart from that, low-aptitude informants with their higher Levenshtein values departed from the high-aptitude and average groups statistically (p=.000).

Figure 6.18 illustrates the distribution of Levenshtein distance values in the aptitude groups by gender.

Mrs McWilliams and the Lightning

When we turned our attention to the second text, an abstract from “Mrs McWilliams and The Lightning” by Mark Twain, we could see at once that the general pattern discovered for “The North Wind and the Sun”, i.e. more variation in values in the English-language samples as opposed to greater relative homogeneity of the German part of the corpus, was replicated. Indeed, the English tokens were much more scattered on the Levenshtein distance scale.

The string length of all English transcriptions was 506 syllables / characters; the corresponding length in German was 492 syllables / characters.

The general cross-group tendencies in the English data were similar as compared to the first text under analysis.

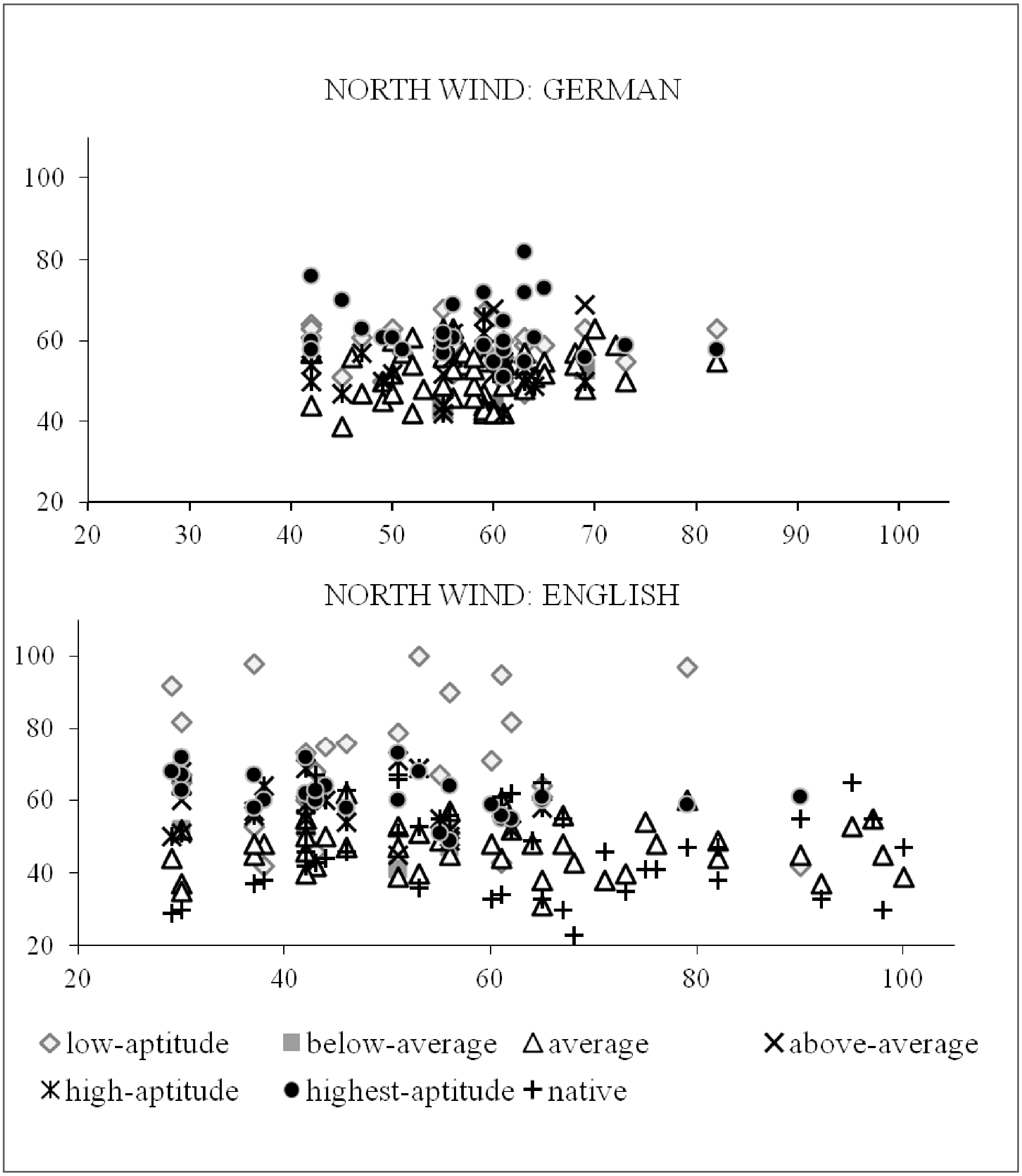


Figure 6.17: Distribution of Levenshtein distance values in "The North Wind and the Sun" in all aptitude groups in the German and English parts of the experimental corpus. On the y-axes: Levenshtein values

More specifically, the greatest degree of inter-speaker agreement in the intonational interpretation of the given text abstract was once again found within the native test persons: about half of their values represented the relative minimum in this condition (~90-110).

However, the other half of the tokens was located higher on the scale and comparable to the values produced by the L1 German informants (~130-170). The low-aptitude and highest-aptitude groups, on the other hand, demonstrated the largest number of cross-speaker differences (~150-230 for the former; ~150-210 for the latter). The remaining groups occupied a common intermediate space, whereby the average and below-average informants tended towards lower values (~130-150) and high-aptitude subjects – towards higher ones (~140-180).

In the German productions of "Mrs McWilliams and the Lightning" no clear separation of groups could be registered. Instead, all values were found within more or less the same range (~110-180). Therefore, the highest- and low-aptitude speakers could no longer be identified by higher values in contrast to the English version of the text.

According to the ANOVA test, the cross-group differences were significant in both languages under investigation (p=.000). The post-hoc Scheffé tests for German did not reveal any consistent statistical trends. Rather, the significant contrasts were sporadic: the highest-aptitude versus the high-aptitude, above-average and low-aptitude speakers (p<.01); the average group opposed to the high-aptitude, above-average and low-aptitude test persons (p<.05). For the English part of the corpus, the Scheffé results were more consistent with the empirical observations. The native group with their generally lower and the low-aptitude subjects with their higher values were statistically segregated from the majority of the remaining speakers (p<.02), except for the highest-aptitude group being statistically similar to the low-aptitude subjects (p=.573) and native speakers – to the below-average group (p=.239). Three other contrasts were found to be statistically significant: between the highest-aptitude speakers, on the one hand, and the average and below-average groups, on the other (p<.01), and between the high-aptitude and average informants (p=.002).

Figure 6.19 is a graphical representation of the distribution of Levenshtein distance values in the aptitude groups regardless of the gender factor.

The division of data by gender did not bring any unexpected contradictions with the general-corpus tendencies as well as the trends stipulated for "The North Wind and the Sun".

All Levenshtein distance values in German female speech were intertwined within the same range (~110-170), and there was only slightly more concentration of tokens around the mean as realized by the high- and highest-aptitude groups. The male productions in the German language, on the contrary, allowed of a clear separation of the average-subjects' sample from the rest of the tokens. The average speakers appeared to have more agreement in prosodic text interpretation evidenced by lower Levenshtein values (~120-150 vs. ~150-170 in the other groups). A similar, though less pronounced, tendency was discovered for the previous text as well.

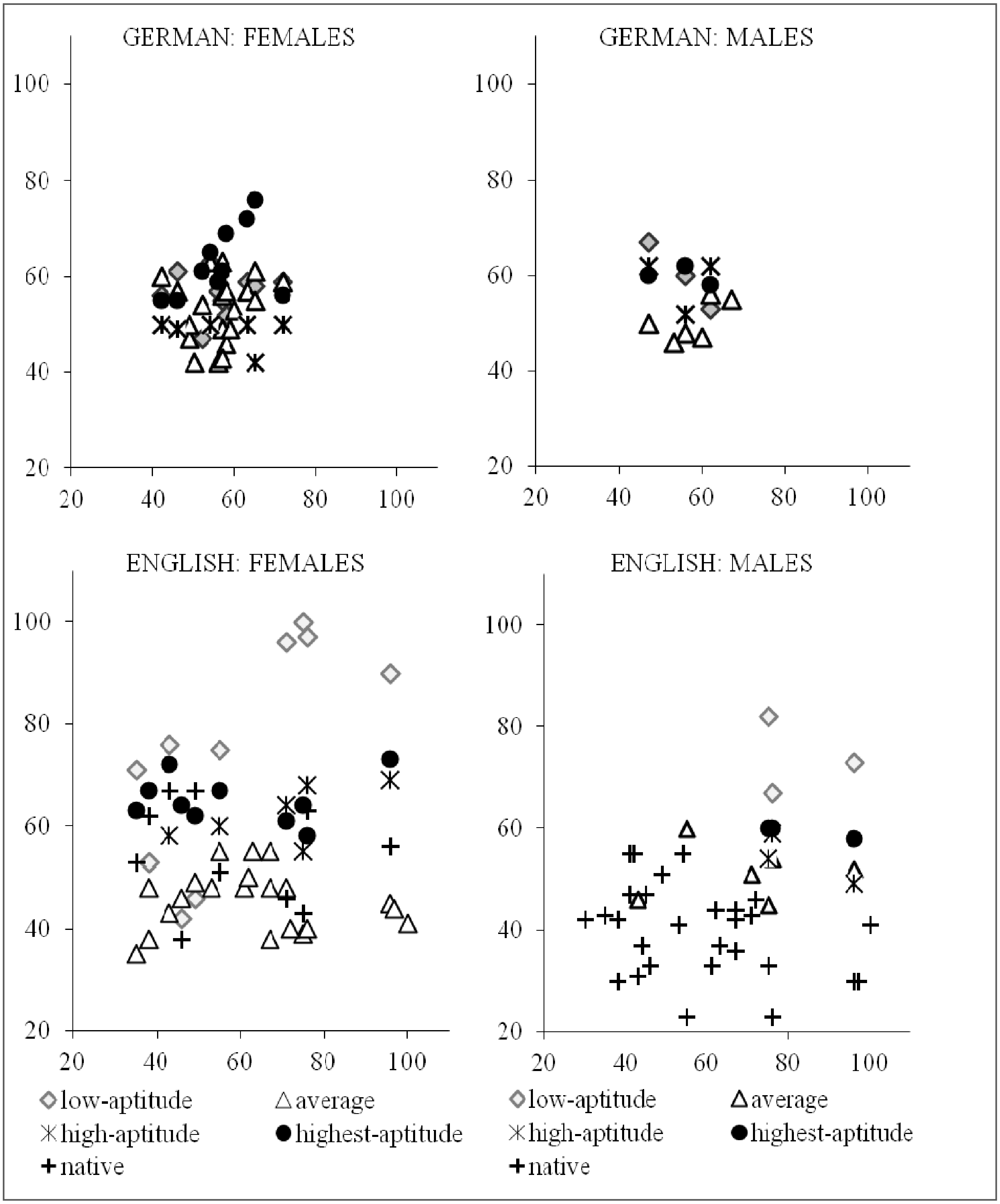


Figure 6.18: Distribution of Levenshtein distance values in "The North Wind and the Sun" in all aptitude groups by gender in the German and English parts of the experimental corpus. On the y-axes: Levenshtein values

The distribution of tokens in the English-language productions also reminded of that proper to "The North Wind and the Sun".

The native and average female speakers once again showed affinity with regard to the text interpretation, i.e. by yielding similar-range Levenshtein values (~120-160), attesting to greater agreement in these two groups. The high-aptitude female group appeared to follow the highest-aptitude group performance in this condition. However, there was slightly more agreement between the former subjects. The highest Levenshtein distance values were demonstrated by the highest-aptitude (~160-210) and especially low-aptitude (~160-230) informants.

The pattern in the English-language male utterances was not that straightforward. The native-speaker realizations could thereby be split into two zones on the Levenshtein scale: one clearly separated from the general value cloud (~90-110) and the other one merged with the average-group tokens (~130-150). The remaining values lay higher than the ones just described (~150-190) and suggested no other trends.

The cross-group differences between the groups were significant both for female and male speech ($p<.01$). The significant cross-group comparisons in the Scheffé tests in German were by far not numerous. Thus, no contrast yielded *p*-values below the significance level in the German female speech. For the male data in German, two tests reached the significance level ($p=.000$), namely, those between the high-aptitude and average groups and between the average and low-aptitude speakers.

The English realizations only attested to one statistical trend as a result of the Scheffé procedure: native male subjects consistently differed from the remaining groups' performance ($p=.000$). The other significant results did not seem to be systematic. In male speech these were the oppositions of the low-aptitude and average groups ($p=.000$) and average and high-aptitude subjects ($p=.027$). In the female data, significant differences were registered for the low-aptitude group versus the high-aptitude, average and low-aptitude informants ($p=.000$); and the highest-aptitude test persons' sample compared with the average and native-speaker productions ($p=.000$).

Figure 6.20 presents the distribution of Leveshtein distance values in the interpetations of "Mrs McWilliams and the Lightning" in the all aptitude groups influenced by the gender factor; Appendix Tables 29, 30 reflect all the results of the post-hoc Scheffé tests.

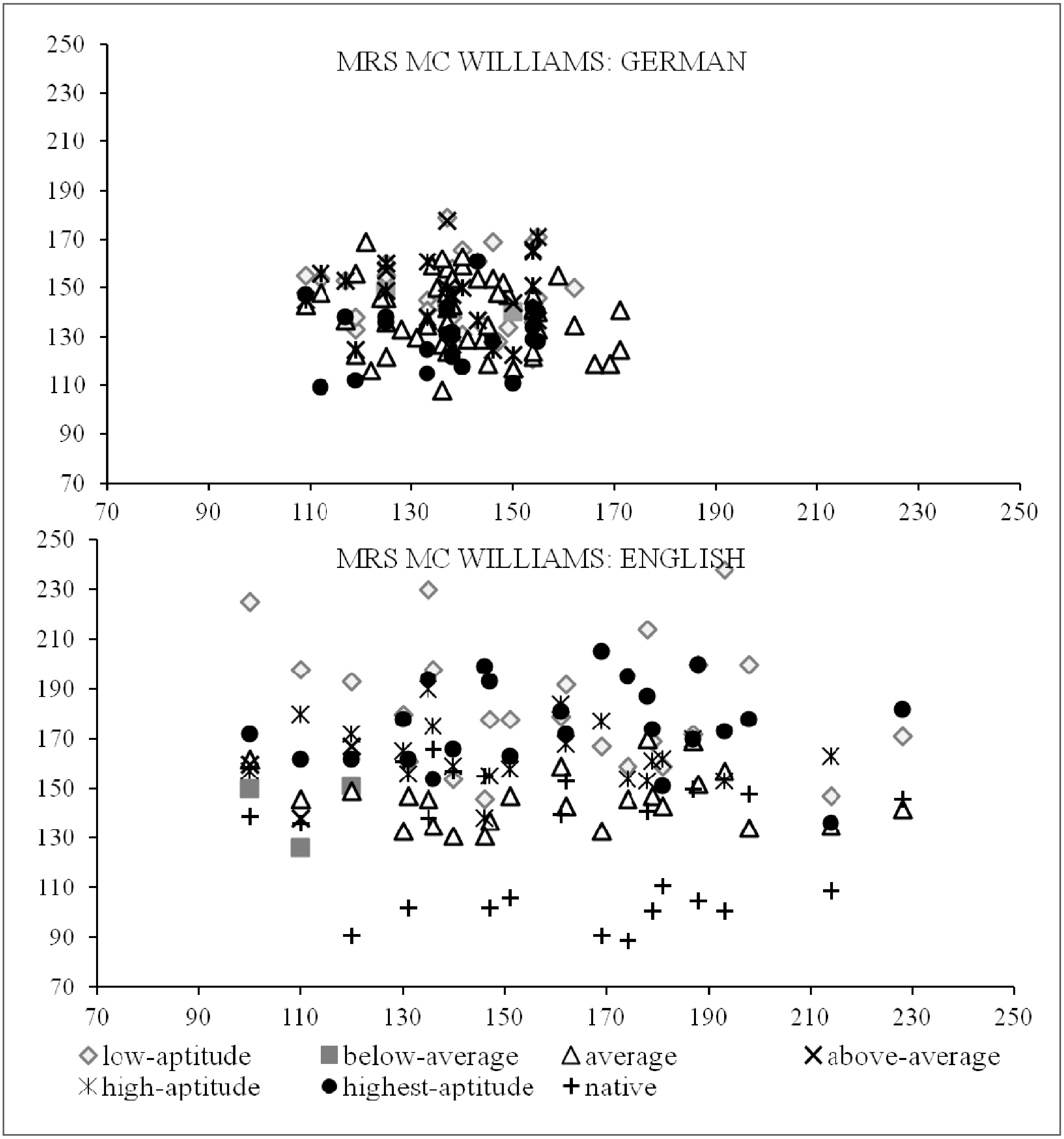


Figure 6.19: Distribution of Levenshtein distance values in the abstract from "Mrs McWilliams and the Lightning" in all aptitude groups in the German and English parts of the experimental corpus. On the y-axes: Levenshtein values

6.3 Summary

In Chapter 6 we dwelled upon the distribution of phonological categories in the current experimental corpus. Such an investigation included two stages: first we analyzed the frequencies of occurrence of ToBI pitch accents and boundary tones; secondly, intonational transcriptions were compared in an attempt to discover the possible correlation between the relative levels of uniformity / variability in prosodic text interpretation and pronunciation aptitude.

With respect to the tendencies in the distribution of ToBI pitch accents and boundary tones, we first looked at the distributional peculiarities of phonological categories in all of the investigated aptitude groups with respect to the gender and speech type factors, as well as in the whole corpus in general. The data was analyzed by language, English and German, in order to identify the presence / absence of cross-linguistic phenomena. Structurally, frequent pitch accents and boundary tones were the focal point of the given study.

Pitch accents

One of the major tendencies regarding the distribution of **rising tones L*H** in the current experimental corpus was related to the L1 / L2 behaviour of the two higher-aptitude groups, who appeared to suppress their native-language tendency in the employment of this category: in German L*H accents have no context restrictions, whereas in English their distribution is rather limited. We could therefore observe lower percentages of rising tones produced by the highest-aptitude and high-aptitude groups in their L2 speech in the following conditions: during the general corpus analysis, in read speech realizations; in female productions without regard to speech type and in female read and spontaneous speech. Notably, such performance was closest to the native English speakers. The other groups employed rising tones more extensively in the above-mentioned conditions, with the exception of the below-average male subject who incidentally produced rising tones with a similar frequency of occurrence as the native male subjects.

The distribution of **simple high targets H*** appeared to suggest one consistent pattern with respect to the native subjects' performance: they generally produced greater percentages of H* accents than the L1 German informants. This trend was supported by the general-corpus investigation and in samples split by speech type. With the application of the gender factor, we could see that male speakers were mainly responsible for such a separation of native and non-native test persons, both with and without regard to the speech type factor. The only other group departing from the generally homogeneous distribution patterns in all conditions was the below-average informant, who applied the category in question more extensively than the remaining male subjects in male speech in general and in English spontaneous productions.

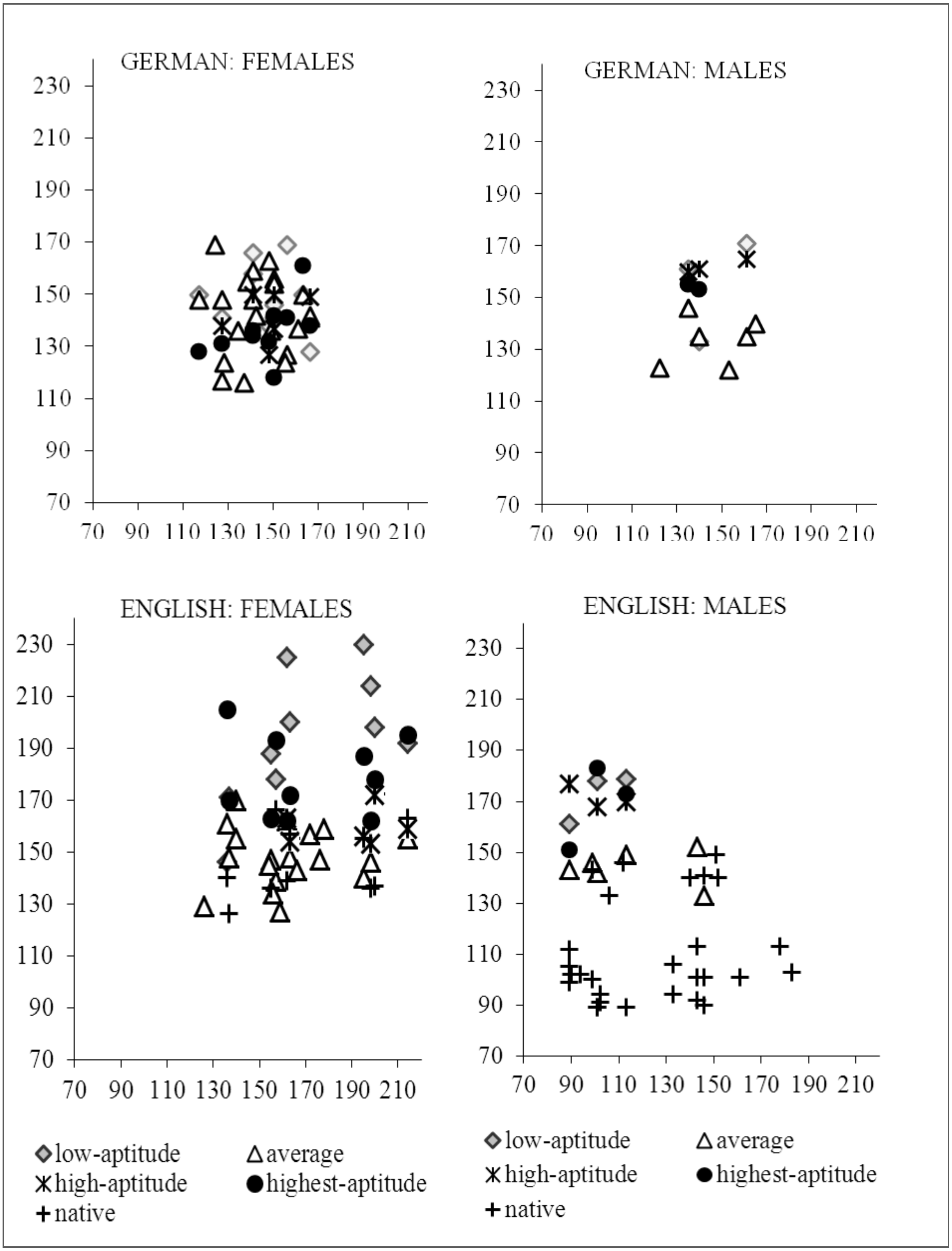


Figure 6.20: Distribution of Levenshtein distance values in the abstract from "Mrs McWilliams and the Lightning" in all aptitude groups by gender in the German and English parts of the experimental corpus. On the y-axes: Levenshtein values

The proportion of **falling tones H*L** was comparable in all groups in the experimental corpus in all conditions. Only the female subjects of high and highest pronunciation aptitude deviated from this pattern by employing the given category more extensively of all female samples in general, as well as female read speech. All other cross-group mismatches were not recurrent.

As far as the **rising-falling tones L*HL** were concerned, we could only state their overall lower percentage in English as compared to German. Moreover, in English spontaneous speech the given events could not even be considered a frequent category due to an insufficient number of tokens in it (~3%).

No consistent frequency of occurrence patterns could be identified in the distribution of **simple low targets L***.

Boundary tones

The distribution of **full low boundary tones L-L%** was somewhat reminiscent of the tendencies in H* pitch accents. More specifically, native speakers could be separated from the L1 German groups due to the higher percentages of L-L% events in their productions in the whole corpus, in the data split by speech type. The closest percentages were exhibited by the highest-aptitude, high-aptitude subjects in all of the just mentioned conditions as well as the below-average group during the general corpus analysis.

When we applied the gender factor to the data, it became apparent that native male speakers exerted greater influence on the above general-corpus trend – in all male speech in general and separated by the speech factor. No such discrepancy could be found within the female subjects: the distance between non-native and native speakers was only marginal.

Turning our attention to the **high full boundary category H-H%,** we could see that the high-aptitude, highest-aptitude and native subjects could thus be once again clustered first during the general-corpus investigation and, further on, in read speech without regard to gender due to their generally lower proportion of high full boundary tones than in the remaining groups. For the two higher-aptitude German groups this was also linked to a slight suppression of the native frequency of occurrence pattern in these conditions. Gender-specific analysis revealed the fact that mainly female speakers were responsible for the trend, both without regard to the speech type factor and in female read speech.

The most conspicuous fact regarding the distribution of **default boundary events %** was their presence in native realizations, since traditionally this category is not part of the English ToBI paradigm. Predictably, the percentage of % tones in L1 German groups was higher in both languages. The application of the speech type factor revealed generally greater percentages of default boundaries in spontaneous speech, as opposed to read utterances. In both speech types, native speakers employed the category in question less extensively than the German groups. Such a separation of native speakers was only supported by male data in English read and spontaneous speech; in the female part of the corpus no such tendency could be discovered.

With respect to the **low and high intermediate tones L- and H-**, no recurrent cross-group patterns could be discovered.

Chapter 7

Investigation of phonetic variation patterns

In Chapter 6 we discussed the variation patterns of ToBI categories, or, in other words, the regularities on the phonological level. However, it is the phonetics of speech that yields insights into the actual realization of these higher-order categories. We therefore attempted such an in-depth analysis of the phonological units, whose distribution and frequency of occurrence peculiarities across subject groups, languages and speech types were described in the previous chapter.

This investigation, in its turn, entailed the analysis of the six *F0* curve parameters produced by the PaIntE model, which we introduced in Chapters 4 and 5. More specifically, these parameters reflect the amplitudes of the rising and the falling sigmoid of an individual *F0* curve (*c1* and *c2*, respectively), the velocity of the rise (*a1*) and fall (*a2*), the alignment of the pitch peak (*b*) and the size of the whole *F0* excursion (*d*). Taken together, all six parameters provide very detailed information about each separate *F0* target. However, it appears that some parameters play a greater role in the definition of specific categories than the others. For example, for a rising tone the parameters concerned with the rising sigmoid would be of more importance than those describing the fall. The latter would, in turn, be more pertinent in the examination of falling tones. Nevertheless the alignment of peaks (parameter *b*) and the total amplitudes of *F0* curves (parameter *d*) seem to be important characteristics of all categories.

Following this logic, we will attempt to describe the realization of ToBI categories, emphasizing the defining parameters for each tonal event.

7.1 Tendencies in the realization of ToBI pitch accents and boundary tones

When analyzing the phonetic realization of ToBI categories as reflected by the PaIntE parameters, we followed a similar procedure as the one applied to the investigation of the phonological regularities, i.e. pitch accents and boundary tones were treated separately.

In each of the statistical comparisons, we looked at the distribution of real values, which give an idea of the actual pitch levels and pitch ranges in the speech samples of each subject group, alongside with rendering the peculiarities in the degrees of *F0* variability. Arguably, such real values may in fact be influenced by the speakers' individual variation. For that reason, prior to undertaking the statistical tests, which will be described further in the current chapter, we standardized the raw data points with the corresponding *z*-scores. Following that procedure, we ran a series of ANOVA comparisons on both the non-standardized and normalized data. The results of these contrasts were very much in line with each other. Consequently, we proceeded with an in-depth investigation of the non-normalized PaIntE parameter values, yielding a more all-round idea of the distribution of these events than *z*-scores. As mentioned above, real targets provide *F0* level, pitch range and variability information, whereby the *z*-scores are only capable of reflecting the degrees of variation. We preserved the findings in the z-scores in the form of box plots for all the investigated categories and conditions. Those can be found in the Appendix to the current investigation.

The infrequent categories remained beyond the scope of the current investigation due to the insufficient number of tokens in them for a valid statistical analysis.

7.1.1 Pitch accents

7.1.1.1 Rising tones: L*H

As suggested by the distribution patterns described in Chapter 6, L*H was one of the most frequent categories in the corpus. Additionally, it appeared to be one of the defining accents from the SLA perspective: it is very common for German, both in the non-final intermediate / intonation phrases and utterance-final, but in English its use in the former condition is rather restricted and less frequent. So we were interested in the L2 subjects' performance in that sense: whether they would retain the phonetic pattern proper to their L1 speech or whether they would adapt to English; and how close different groups would be to the native speakers. The phonological analysis evidenced an accommodation pattern in L*H accents in the high-aptitude and highest-aptitude groups, especially characteristic of the female test persons in read speech. This trend was also seen to be an additional test objective on the phonetic level, where it had to be checked in order to see whether it would be preserved or abandoned in the realization of the fine-grained PaIntE parameters.

In the investigation of the L*H accents, we mainly focused on the variation patterns of parameter *d,* yielding values, which reflect the *F0* levels and excursions of individual pitch curves. Its production in the corpus will be treated in detail in the current section. *c1* (the pitch excursion of the rising sigmoid) was predicted to follow the general pattern set by the amplitude of the whole *F0* excursion in a curve (parameter *d)*, as the rising element takes up a substantial part of any rising tone. Therefore, the description of the statistical results, pertaining to the *c1* parameter, was meant to test this prediction. We will also present a summary of the variation patterns in the alignment *F0* peaks (parameter *b*) and the velocity of the rises (parameter *a1*). Earlier research on these phenomena has been scarce, if at all available. Therefore, any implicit trends in these parameters would facilitate our understanding of the phenomena in question, which, in turn, would contribute to the general phonetic description of such events.

Parameter *d*

We started our investigation by analyzing the realization of the *d* parameter, i.e. the total amplitude of a given *F0* curve.

General corpus analysis

The very first examination of values in all the groups showed that the L1 German subjects produced L*H events with almost equal degree of variability in the two languages under investigation, which implies that the speakers applied a similar phonetic strategy both in the English and German realizations.

Comparing the results across the groups, we discovered the greatest degree of pitch variability within the highest-aptitude subjects in English and German to an equal extent. In that respect, high-aptitude and average subjects also exhibited substantial variability, though lower in degree as compared to the highest-aptitude informants. Lowest was the degree of variability in the low-aptitude, above-average and below-average groups (see Figure 7.1 for the illustration of the above tendencies).

When we looked at the realizations of native speakers, who only produced the test utterances in their mother tongue, it became evident that their variability pattern most closely approached that of the highest-aptitude group, i.e. the L1 English speakers also displayed a substantial degree of variation in the *d* parameter. High-aptitude and average speakers produced a generally comparable, though a more concentrated variation pattern.

It is also worthy of mentioning that the mean values of most of the German groups in the English realizations, except for the above-average speakers, lay above the corresponding value in the native-speaker group. For most L1 German groups the data points lay between ~200-230 Hz, whereby the native speakers produced L*H targets at a lower *F0* level of ~180 Hz. It appeared to be a language-specific phenomenon in the given condition, as the respective values in German were on a similar pitch level as the English-language realizations by the German speakers. The only exception to this trend were the productions of the above-average subjects, whose mean *F0* in both languages was below that in the native English group (~150-170 Hz).

To validate the empirical observations, we compared the group samples statistically by means of running a series of ANOVAs, splitting the data by language. Both in German and English, the differences between the groups were statistically significant according to the standard test (p=.000). The post-hoc Scheffé comparisons presented a more detailed insight into the inter-group differences. Thus, almost all samples in the two languages were statistically different from one another, though with a few non-systematic exceptions.

Figure 7.1 represents the distribution of parameter *d* values in L*H accents in the whole experimental corpus; Tables 1-2 of the Appendix render all the pertinent statistical results of the Scheffé tests in the given experimental condition.

Gender-specific analysis

The next stage of data analysis consisted in investigating the data genderwise, to see if the tendencies proper to the whole corpus would be just as evenly represented in the male and female realizations.

As mentioned earlier in the section, the German groups appeared to preserve the variability patterns across the two languages, as is evident from the general corpus evaluation. This tendency was to a substantial extent supported by the male realizations, excluding the highest-aptitude group, who increased the degree of variability in English, and the above-average group, who, vice versa, produced English utterances in a less variable way. There were also two exceptions to this tendency within the female subjects: the high-aptitude test persons exhibited more variation in English than in German, whereas the above-average subject performed in an opposite fashion in her L2 speech.

The general-corpus peculiarities of pitch level variability in all the subject groups appeared to find more evidence in the female part of the data.

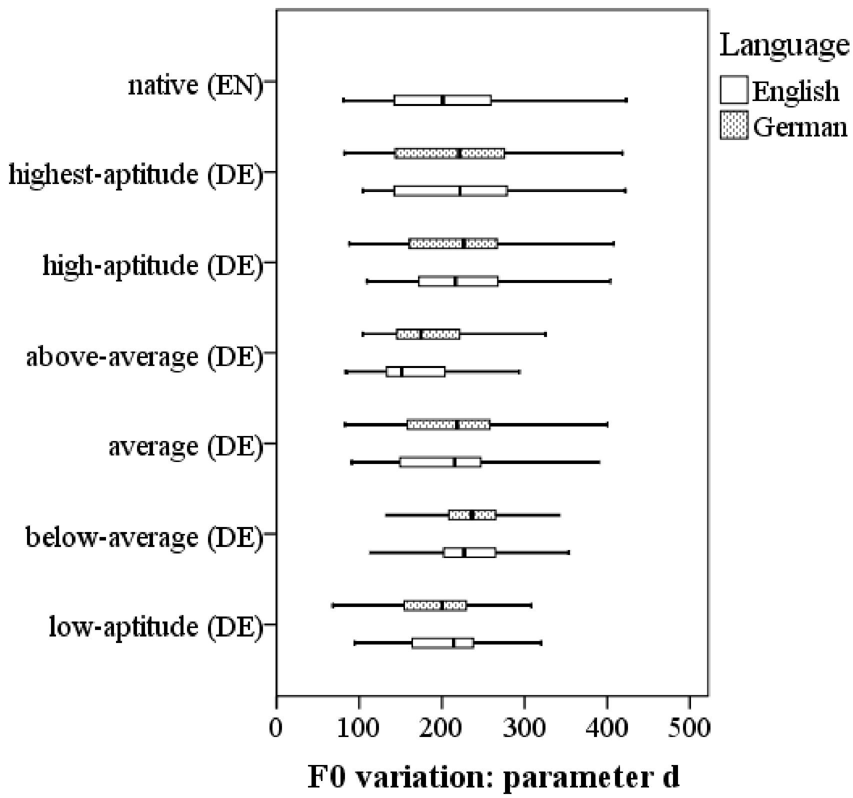


*Figure 7.1: d parameter variation of L*H accents with regard to the F0 level in Hz in all the aptitude groups in English and German*

Thus, in the same way as in the whole corpus, native, highest-aptitude and high-aptitude, as well as, surprisingly, below-average female group realizations were marked by the highest degree of variation in English.

The same was, but partially, true for these groups in German, where high-aptitude speakers exhibited less variability than in their L2 realizations. The degree of variation in English and, for the most part, German was therefore lowest in the average, above-average and low-aptitude groups.

The general tendency of German speakers to realize their utterances on a higher mean pitch level in English than the native speakers was not supported by the female data. Here, the highest-aptitude and high-aptitude speakers clearly hit higher *F0* targets than the remaining groups (~250 Hz). The mean values of the native group were only slightly higher than the remaining German groups (~225 Hz vs. ~220 Hz, respectively), but located on a lower level than the two groups of higher pronunciation aptitude we have just mentioned.

One other peculiarity of the high-aptitude and highest-aptitude subjects was the fact that they raised their mean *F0* when speaking English by about 10 Hz. Quite surprisingly, the same was also proper to the low-aptitude group, whereas for the rest of the subjects there was no such difference cross-linguistically.

These observations were supported by ANOVA and the post-hoc Scheffé tests. The former seconded the fact that the inter-group differences were generally significant (p=.000). Most Scheffé comparisons in English and German as realized by the female informants also yielded significant results (p<.031). Non-significant exceptions did not appear to be systematic.

The male data demonstrated less consistency with the trends discovered for the entire corpus.

The degree of pitch level variation in all groups was almost equal, contrary to the whole-corpus predictions and observations of the female part of the database. The only exception was represented by the below-average speaker, who displayed the smallest degree of variability in L*H accents. The variation ratios in English and German were also not homogeneous as realized by L1 German speakers. For instance, highest-aptitude and average speakers expanded their pitch range in English as compared to the L1 samples, whereas the above-average speakers reduced it. In the remaining groups the degrees of variation were almost equal across the languages.

With regard to the mean pitch level, it should be noted that in German it was slightly highest in the low-aptitude and high-aptitude, as well as above-average groups (~150-160 Hz), and lowest in the productions of the below-average male speaker (~120 Hz). In the English part of the corpus, high-aptitude speakers clearly demonstrated a higher overall mean *F0* than all the other groups (~170 Hz vs. ~150 Hz, accordingly). The below-average male speaker once again produced conspicuously low mean values (~130 Hz).

The cross-group differences in the male groups were significant in both languages, according to the ANOVA outcome (p=.000). Most inter-group mismatches were also statistically significant (p<.05), with more pitch level overlap between the groups in English than in German.

Figure 7.2 illustrates all the observations of pitch level and variability peculiarities of the d parameter in L*H accents in male and female realizations across the aptitude groups; Tables 1-2 of the Appendix render the p-values of all Scheffé contrasts.

Corpus analysis by speech type

After applying the gender factor to the data, we also investigated the corpus by speech type. As a result, generally similar tendencies in read and spontaneous realizations in English and German were discovered, as realized by both the native and the L2 speakers.

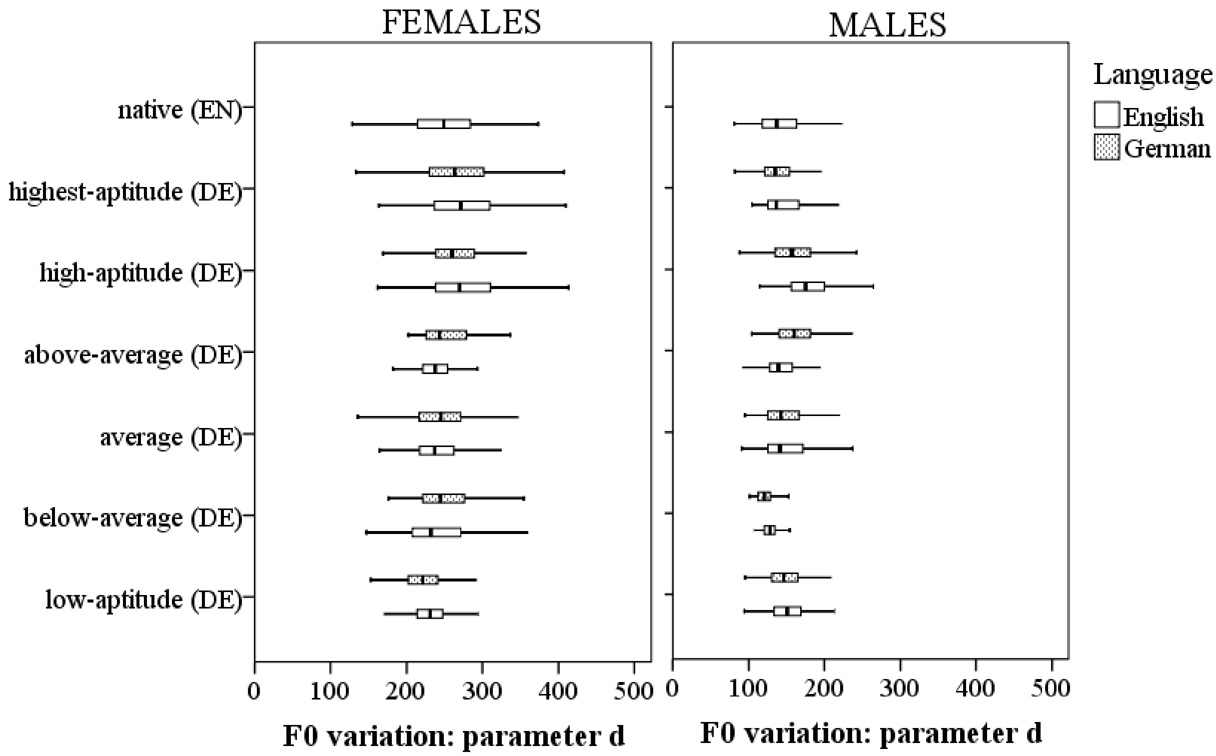


*Figure 7.2: d parameter variation of L*H accents with regard to the F0 level in Hz in all the aptitude groups in English and German genderwise*

Thus, the greatest degree of variability in pitch levels could be found in the highest-aptitude group under all speech type and language conditions.

The native-group pattern in the two speech types was closest to that observed for the highest-aptitude subjects. In read speech, these two groups were closely followed by the high-aptitude and average speakers. However, in spontaneous realizations this was only characteristic of the average group: the high-aptitude test persons followed the lower variability pattern proper to the remaining groups in this respect.

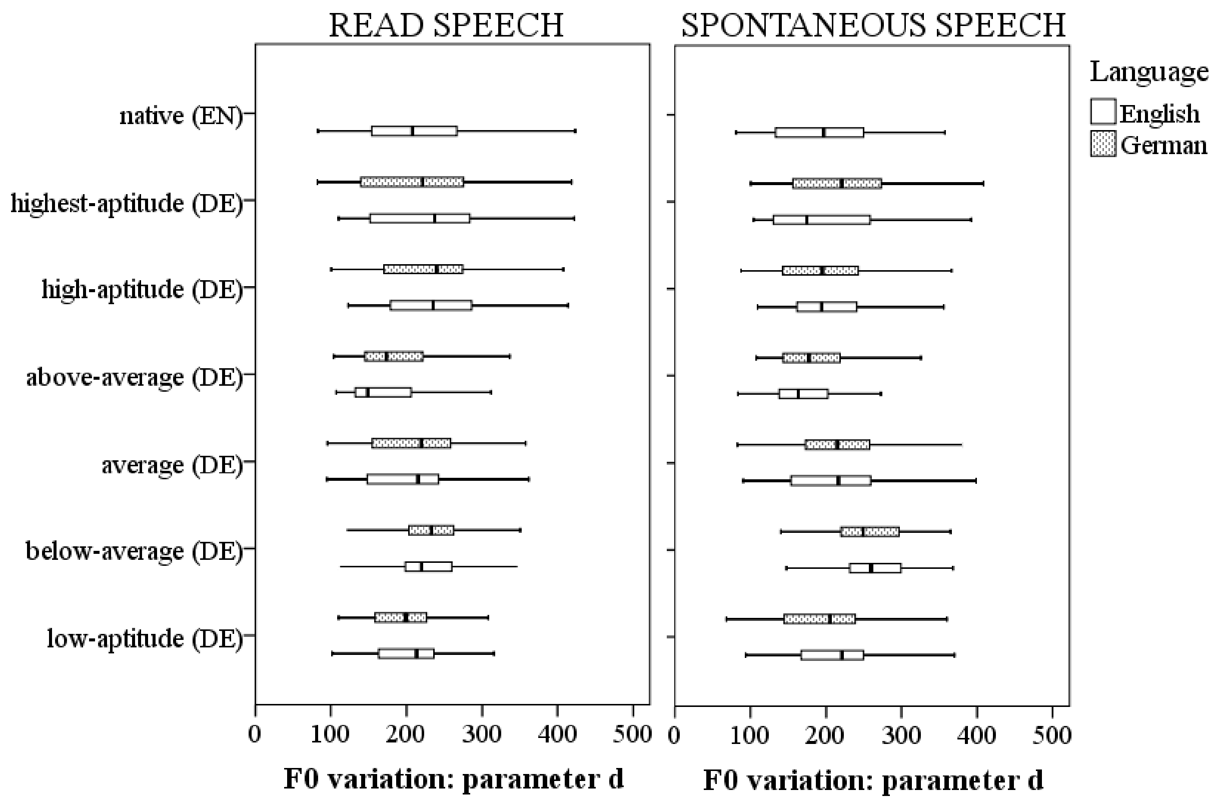


*Figure 7.3: d parameter variation of L*H accents with regard to the F0 level in Hz in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

One of the tendencies discovered for the whole corpus, i.e. the fact that L1 German speakers tended to produce utterances on a higher mean pitch level than the L1 English speakers, was supported by the read speech data. All the German groups, except for the above-average subjects (mean *F0* level of ~140 Hz in English and ~170 Hz in German), had higher mean *F0* values in both languages (~210-240 Hz) than the native speakers (~190 Hz). And it was most evident in the high-aptitude and highest-aptitude group samples in English, exhibiting the relative maximum values.

Spontaneously produced utterances did not reflect this trend just as uniformly. In this part of the corpus, above-average test persons also realized the test material on a generally lower *F0* level than the other groups in both languages at issue (~170-180 Hz vs. ~200-250 Hz). However, in English these speakers were approached in performance by the highest-aptitude group, producing *d* parameter tokens on the same low pitch, though with a much greater degree of variability. Still another unexpected observation concerned the outstandingly high mean values within the below-average subjects (~250-260 Hz in both English and German).

ANOVA tests demonstrated significant differences between the groups in all speech-type and language contrasts (p=.000). Most post-hoc tests, with only a few exceptional contrasts, resulted in cross-group statistical significance (p<.05).

Figure 7.3 and Tables 1-2 of the Appendix contain graphical and statistical evidence for the description of *d* parameter value distribution in L*H accents by speech type.

Corpus analysis by speech type and gender

When we applied both the speech type and the gender factor to the data, it became evident that one of the previously discovered tendencies, i.e. the greater degree of variability in the native, high-aptitude and highest-aptitude groups as compared to the other informants, was most characteristic of female test persons in English read speech. Here, the contrast between these three groups and the other speakers was quite apparent, although average speakers also exhibited substantial variability, slightly lower than that produced by native subjects.

The German-language read-speech realizations of female speakers presented a conspicuous variation pattern, where only the highest-aptitude group could be singled out for their greater degree of variation of the *d* parameter in L*H accents, as compared to the remaining subject pool. The high-aptitude and low-aptitude test persons, vice versa, exhibited the least variable samples.

The second overall trend, i.e. the raising of pitch level by the German subjects was not indisputably supported by the female read speech data. Rather, it was the cluster of the native, highest-aptitude and high-aptitude speakers, who produced the experimental utterances on a higher mean pitch (native speakers: ~235 Hz; high-aptitude and highest-aptitude: ~250 Hz; other groups: ~220 Hz) in both the English and the German utterances.

The inter-group differences in female read speech appeared to be significant as evidenced by the ANOVA results for all the experimental conditions (p=.000). Both in English and German, highest-aptitude and high-aptitude subjects revealed statistical similarity (p>.67), but different from the remaining group samples (p<.05). One other instance of statistical dissimilarity was found in German for the mean values realized by the low-aptitude group, which were generally lower, hence a statistical difference from the other test persons (p<.05).

In spontaneous speech, female productions were marked by comparable degrees of variation in both languages, with the exception of the English and German samples of the low-aptitude group and the realizations of the above-average speaker in English, whose values were more uniform.

As far as the pitch level is concerned, in German it lay in the interval between ~240-250 Hz in all the groups, except for the low-aptitude informants with a slightly lower mean pitch (~230 Hz).

Unexpectedly, the highest values in English were produced by the below-average informants (~260 Hz), whose performance was followed by the subjects of highest and high aptitude (~250 Hz).The remaining groups produced lower mean *F0* values (~230-240 Hz).

According to ANOVA's, the mean *F0* values were statistically different across the groups in the two languages (p=.000). However, there was less heterogeneity in the data overall, i.e. most cross-group comparisons in the post-hoc tests were not significant. Only the low-aptitude test persons exhibited a consistently different performance, most probably due to their tendency for lowest *F0* values in both English and German (p<.02). The average subjects, however, were statistically similar to the low-aptitude group in this condition (p>.9).

The male part of the corpus contained less homogeneity and consistency than the corresponding female samples in both speech types. However, one of the two general trends discovered for the whole corpus, i.e. the raising of the mean *F0* by German speakers, was supported by almost all the spontaneous speech samples in English, excluding the highest-aptitude realizations. The latter performed in a similar way as the native group in this respect (~130-160 Hz in the non-native speech, with the highest values in the high-aptitude group, vs. ~120 Hz in native and highest-aptitude groups). In English read speech, the average pitch level was almost equal for the native, highest-aptitude and above-average informants (~130 Hz). Low-aptitude group tokens exhibited a higher mean *F0* (~150 Hz).

Finally, the highest average pitch level was found in the high-aptitude group (~170 Hz), similarly to their performance in spontaneous speech. The below-average subject once again performed idiosyncratically by producing lower general values than all the other groups (~120 Hz).

Statistically, the inter-group differences in the English part of the corpus in male speech were significant as stated by the general ANOVAs (p=.000), and so were most of the post-hoc Scheffé tests, with just a few non-systematic exceptions. Cross-group comparisons in spontaneous speech were also, for the most part, statistically significant.

When we turned our attention to the German-language productions, we found some similarities as well as deviations from the English part of the corpus – within each group and across the subjects of varying pronunciation aptitude. Thus, in read speech average and low-aptitude speakers retained the same mean *F0* level as the one revealed for their L2 English speech. A lowering of average pitch was characteristic of the highest-aptitude, high-aptitude and below-average groups by ~10-20 Hz; above-average speakers, on the contrary, raised their mean *F0* level in their native language by ~20 Hz. With respect to the *F0* level in general, the highest-aptitude and low-aptitude groups were marked by the lowest mean *F0* values in German read speech (~120 Hz). In all the other groups under investigation the average pitch levels lay at ~140-150 Hz.

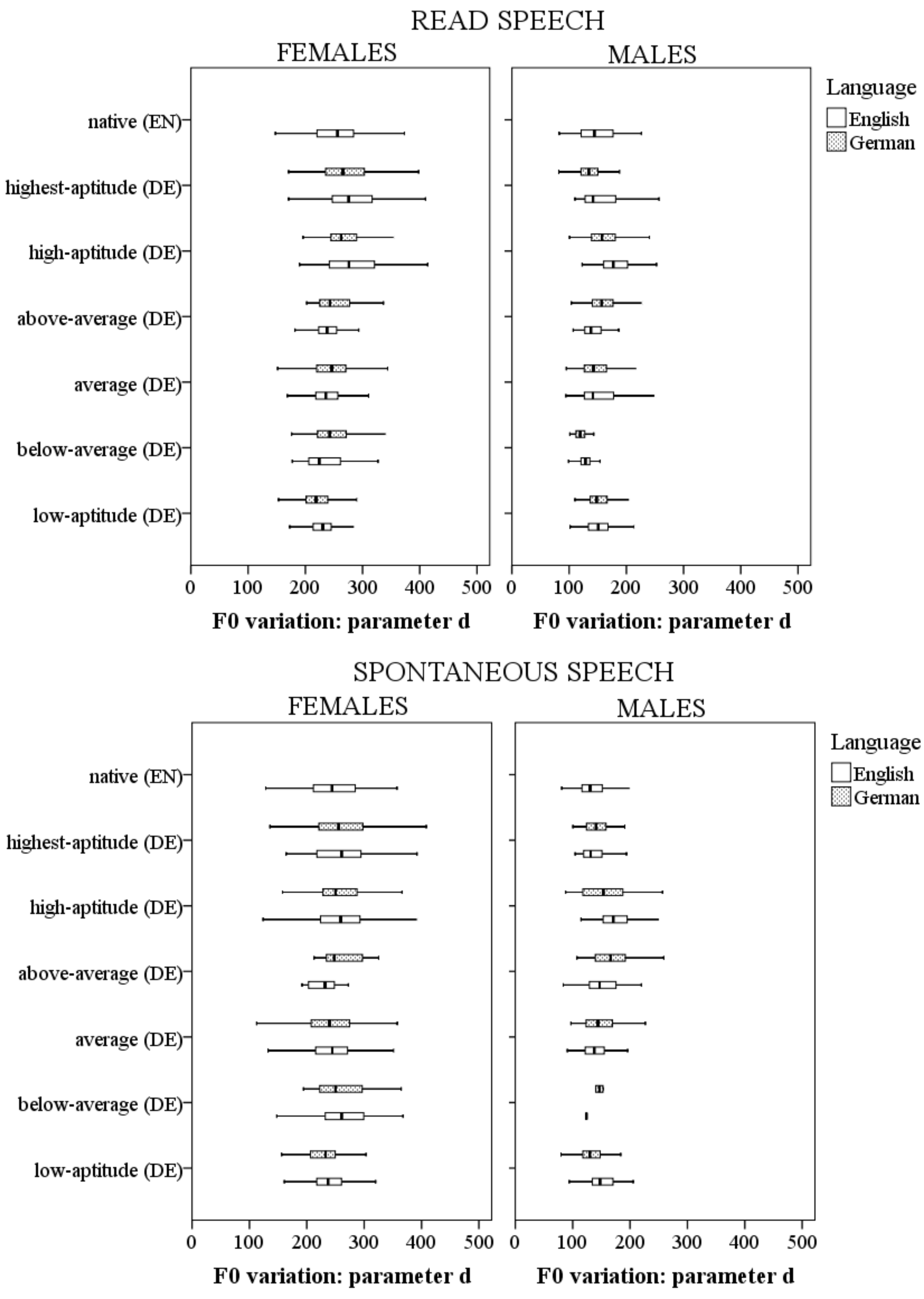


*Figure 7.4: d parameter variation in L*H accents with regard to the F0 level in Hz in all the aptitude groups in English and German read and spontaneous speech by gender – male versus female realizations*

Notably, even such slight differences in the mean *F0* of the latter groups were found to be significant in most Scheffé contrasts ($p<.02$). On the whole, there were only three non-significant comparisons in this experimental condition: the one between the high-aptitude and average groups ($p=1$), as well as the one between the low-aptitude subjects and the average informants ($p=.473$).

Spontaneous speech in German replicated the above finding of the lowest mean pitch found in the highest-aptitude and low-aptitude groups (~120 Hz). In the remaining groups the average pitch was higher (~150-160 Hz), with the above-average subjects exhibiting the maximal values.

Most of these differences appeared to have no statistical significance according to the Scheffé tests ($p>.1$). Nevertheless the above-average group with their higher than average *F0* level and the low-aptitude speakers demonstrating relatively low pitch did differ significantly from some of the samples. In the case of the former subjects these significant comparisons involved the highest-aptitude, high-aptitude and above-average groups ($p=.000$). The low-aptitude speakers differed in their mean *F0* from the high-aptitude, above-average and average groups ($p=.000$).

As far as the degrees of variability are concerned, we first analyzed the English-language samples. Striking in read speech was the similarity between the native-speaker, highest-aptitude and average group male samples. The below-average male speaker, on the other hand, demonstrated still another instance of idiosyncrasy: tokens had far less variability than the other groups. The same observation was also valid for his L1 German speech. The remaining read-speech samples occupied a somewhat intermediate position between these variability extremes.

English spontaneous speech appeared to have more agreement between the groups in the degrees of variation, but for the outstanding performance of the below-average subject mentioned above. Thus, in all groups the degrees of variability were almost equal.

The investigation of German-language spontaneous realizations followed the analysis of English speech. As a result, it became apparent that read speech here was slightly more uniform with respect to the degrees of variability in the male groups than spontaneous utterances. More specifically, the samples realized by the high-aptitude, above-average and average speakers were almost equally variable. The below-average speaker, contrary to that, had the most mean-concentrated values of all the subjects.

In spontaneous speech, the highest-aptitude subjects demonstrated conspicuously more variable tokens than the remaining groups. The above-average and average test persons had slightly less variability in their values; and the least variable samples were to be found within the low-aptitude, highest-aptitude and, most prominently, below-average group realizations.

Figure 7.4 and Tables 1-2 in the Appendix illustrate the distribution of the values in parameter *d* of L*H events in read and spontaneous speech by gender.

Parameter *b*

The second PaIntE parameter under analysis was the alignment of the *F0* peak within an individual L*H target, i.e. parameter *b*. We will only present a summary of its distribution in the subject groups, whereby all the corresponding demonstration materials (box plots and tables of statistical significance based on Scheffé post-hoc tests) can be found in the Appendix to the current investigation.

General corpus analysis

The very first general observation regarding the variation of this parameter in the experimental corpus was the fact that its values were generally more variable in the German-language realizations in all groups. In English, all L1 German speakers demonstrated decreased variability of the *b* parameter: this phenomenon was quite pronounced as realized by most informants, except for the highest-aptitude and high-aptitude speakers. In these two groups only a slight drop in variability could be observed. Overall, the native subjects exhibited the greatest degree of variability in English. Their performance was most closely approached by highest-aptitude and high-aptitude speakers, as well as, unexpectedly, subjects of below-average pronunciation aptitude[18].

Statistically, the differences between the groups were significant in both languages under investigation, as evident from the ANOVA results (p=.000). The post-hoc Scheffé tests established statistical distinctions between the groups in a pairwise manner. In particular, the German experimental data showed that the above-average informants were marked by a statistically different performance as opposed to most of the groups (p<.04), with the exception of the below-average speakers (p=.330). In English, low-aptitude test persons differed statistically from all the groups (p<.05), excluding the average speakers (p=.469). The L1 English subjects, on the other hand, appeared to have produced mean values similar to those in most groups, but for the just mentioned low-aptitude group (p=.000) and the average speakers (p=.000)[19].

Following the same logic as in the investigation of the amplitudes of the *F0* peaks in all groups (parameter *d*), we proceeded by applying the two factors, i.e. gender and speech type, to the data, both individually and in combination.

18 Figure 6 of the Appendix, illustrates the distribution of parameter *b* in L*H accents in the whole experimental corpus.
19 Tables 5-6 of the Appendix render the results of Scheffé tests for parameter *b* in all statistical comparisons pertaining to L*H accents.

Gender-specific analysis

The genderwise comparison demonstrated an almost equal degree of variability in the German realizations of the female speakers. The same was, but partially, true of the male groups in this respect. Indeed, highest-aptitude and high-aptitude, as well as average and low-aptitude speakers aligned *F0* peaks within pitch curves in L*H accents with similar variability. Samples produced by the above-average male informants, on the other hand, were more variable in this condition, whereby the below-average male test person exhibited the smallest degree of variability.

The most interesting observation for the English part of the corpus was related to the performance of male subjects of highest and high aptitude, who replicated the native-speaker pattern. The values of the other male groups were less variable, both compared to the higher-aptitude and native male speakers and to the corresponding German-language tokens. In female realizations the variability ratio was less clear: whereas highest-aptitude, above-average and below-average informants displayed a degree of variation similar to that of the native subjects; high-aptitude, average and low-aptitude test persons could be grouped in the other cluster based on the fact that their samples were less variable with respect to the *b* parameter in English.

ANOVA's attested to statistical significance for the male and female samples in the alignment parameter of L*H accents, both in English and in German ($p<.04$). More specific information, however, could be retrieved from the post-hoc Scheffé tests. Thus, the male subject of below-average aptitude once again displayed exceptional performance in a way that his productions were statistically different from all the other male groups in German ($p<.04$). It should be mentioned here that female below-average speakers also deviated from most of the female informants in this condition ($p<.03$), with the exception of the above-average group ($p=.963$). Still another sample departing from most other speakers were the male subjects of low pronunciation aptitude ($p<.05$), who only agreed with highest-aptitude and below-average test persons ($p>.4$) in statistical terms. The remaining contrasts did not appear to be systematic

For the English data, the above-mentioned observation of a relative subdivision of male speakers, based on their aptitude and proximity to the variability pattern produced by the native speakers, was supported by the Scheffé tests. Here, the native-subject sample was not statistically different from those of the highest-aptitude, high-aptitude and above-average groups ($p>.8$). Scheffé comparisons attested to the male below-average speaker's aberrant performance in English as well, by rendering statistically significant *p*-values in this condition ($p<.005$). Similar to their general variability pattern, female realizations were also less consistent in relation to post-hoc contrasts. The only female group not differing consistently from the rest of the test persons were the highest-aptitude informants ($p>.1$)[20].

[20] Figure 7 and Tables 5-6 of the Appendix are referred for a demonstration of all cross-gender findings.

Corpus analysis by speech type

After comparing the data gender-specifically, we looked at the distribution of the *b* parameter in read and spontaneous speech. The very initial look at the corresponding values allowed of an interesting observation: spontaneous speech samples in both English and German, as well as the German realizations in read speech, were generally more variable than the respective read speech tokens in English (see Figure 8, left-hand charts, Appendix, for a demonstration). This contrast was in fact quite striking, as all the groups, including the native speakers, displayed a degree of variability in English read speech, reduced by around 50%, as compared to the corresponding German-language samples. Of course, it is only possible to speculate about native speakers' performance in German: would the contrast be as distinct as in the L1-L2 productions of the German speakers? Answering this question is beyond the scope of the current investigation. However, additional tests could support or disprove the universality of the trend at issue.

There was no visible difference between the read-speech realizations in English, as far as the degree of variability is concerned. The German read speech was also almost equally variable across the aptitude groups. Spontaneous speech revealed some, but not too evident, differences in variation patterns between the informants. Thus, in German the low-aptitude speakers demonstrated values more concentrated around the mean than the remaining groups. English data, in turn, was marked by a greater degree of variability in the highest-aptitude and below-average groups.

Cross-group differences were registered as significant by most of the ANOVA tests ($p<.003$). Scheffé post-hoc comparisons stated that all spontaneously produced tokens in German were not different from one another statistically ($p>.056$).

In English read-speech samples, most comparisons did not yield significant results ($p>.05$), and those found significant did not reveal any consistent pattern. The same observation was valid for English spontaneous speech[21].

As the final step of analyzing the peculiarities of *F0* peak alignment, evidenced by the distribution of the PaIntE parameter *b*, we split the data both according to the respective subject gender and the speech type.

Corpus analysis by speech type and gender

As a result of the investigation of the data by speech type and gender, we were able to second the observation put forward earlier in this section. More specifically, gender-specific samples in read speech exhibited the same trend as the one detected for the whole corpus: there was less

21 Appendix Figure 8 illustrates the distribution of parameter *b* in read and spontaneous speech; Tables 5-6, in their turn, present the cross-group statistical comparisons.

variability in the alignment of *F0* peaks in English as compared to the German-language data. One exception was, however, represented by the idiosyncratic productions of the above-average female speaker: she displayed more variation than the remaining tokens pertaining to the other groups - male and female. Casting out this exception, the female productions were relatively homogeneous, especially in German.

Male groups demonstrated less homogeneity with regard to the *b* parameter in German read speech. Whereas highest-aptitude and high-aptitude, as well as average and low-aptitude subjects produced comparably variable values, above-average group tokens had a significantly greater and below-average speaker's samples a visibly lower degree of variability in the given experimental condition. English productions were, for the most part, uniform in degrees of variability.

ANOVA disclosed significant inter-group differences in parameter *b* ($p<.01$) in German read speech. According to the post-hoc Scheffé tests, most female-group contrasts were not significant, similarly to male speech in that respect. The few significant results, in their turn, did not put forward any cross-group trends.

To the systematic phenomena in English read speech belonged the above-average female speaker's performance, differing statistically with respect to the alignment of *F0* peaks from all the remaining female groups ($p<.01$). Only below-average group utterances were comparable in the current condition ($p=.130$).

The below-average male subject exhibited idiosyncratic statistical behavior in English read speech, with significant *p*-values in all contrasts ($p<.03$), except when his sample was compared with that produced by the highest-aptitude and low-aptitude groups ($p>.1$).

Finally, examining the spontaneous speech data gender specifically, we made several additional observations on the nature of *F0* peak alignment in the corpus. One of them pointed out to an opposite trend in male vs. female behavior, when compared with the read speech samples. We thus stated earlier that female realizations in read speech were marked by a greater degree of homogeneity than the utterances produced by male speakers. In spontaneous speech, however, it was male speech that demonstrated more uniformity. We ought to point out for this condition that the below-average speaker did not produce a sufficient number of L*H tokens for a valid statistical comparison.

German female spontaneous speech was also relatively homogeneous, as far as the degree of variability is concerned. However, the above-average speaker's mean values were visibly higher than the mean in the remaining groups. In English, the same subject was identified by the lowest degree of variability in the alignment parameter. Other groups produced clearly more scattered *b* parameter values, with the greatest degrees of variability in the highest-aptitude group, closely followed by the native, high-aptitude and below-average speakers.

Once again, ANOVA's yielded significant differences across the groups (*p*<.05). No statistically significant contrasts were found in male groups in German spontaneous speech (*p*>.1). Most female groups, however, differed in this condition from the above-average speaker (*p*<.04) The remaining tests did not point to any systematic trends[22].

Parameter *a1*

At the final stage of the investigation of L*H accents we analyzed the two PaIntE parameters specifically describing the rising sigmoid of an individual rising tone token. We therefore first looked at the parameter *a1,* which defines the steepness of the rise, i.e. specifies how fast the *F0* peak is reached within an intonational curve.

Following the same order set for parameters *d* and *b*, we initially explored the peculiarities of parameter *a1* functioning in the whole corpus. As a result, no significant deviations between the groups in the degree of variability were found, neither in the English nor in the German data. Only the below-average speakers demonstrated more variation in English than the remaining subjects.

Despite this relative homogeneity with respect to variability, ANOVAs still attested to statistical significance in sample comparison (*p*=.000). However, in German, only above-average and high-aptitude speakers exhibited contrasting performance: the former group – as opposed to highest-aptitude and high-aptitude informants (*p*<.04); the latter – differing from above-average, average and low-aptitude test persons (*p*<.03)[23].

Splitting the data by gender did not reveal any striking disparities between the groups in degrees of variability. The male samples were indeed quite uniform in this sense. In female productions, the only dissimilarities were found in the performance of the above-average speaker in German, whose tokens were visibly less variable than those produced by the other groups.

ANOVA's stated a statistical significance of cross-group differences (*p*<.03). However, a more detailed analysis using Scheffé tests pointed out to no significance in all contrasts between the male groups in German. In female samples of this language statistically significant results were also not numerous. In the English data no groups could be singled for consistent deviation from the other speakers with respect to parameter *a1* of L*H accents[24].

---

22 Refer to Figures 9-10 and Tables 5-6, Appendix, for a detailed insight into the above findings of gender-specific samples in read and spontaneous speech of parameter *b* of L*H accents.
23 Tables 3-4 in the Appendix reflect the statistical Scheffé findings in the distribution of the *a1 parameter* in the whole corpus, and Figure 1 graphically presents the variation of *a1* values in this condition.
24 Tables 3-4 and Figure 2 of the Appendix are to be referred for cross-gender peculiarities in the realization of *a1* tokens in L*H accents.

When we applied the speech type factor to the data, we could see that read speech samples were mostly homogeneous, as far as the degree of variability is concerned. The only observable differences here concerned the below-average speakers' performance in English, whose *a1* parameter values were also more variable than in the remaining groups.

Contrasts in spontaneous speech were, however, more pronounced. More specifically, there was distinctly less variability in the above-average group realizations. The same was proper to the low-aptitude speakers and the below-average informants in English. The degrees of variability in both languages were comparable as produced by the remaining informants.

Mean values of the *a1* parameter were statistically different from one another, according to the ANOVA's ($p<.03$). Scheffé tests, in their turn, only specified several statistically valid cross-group contrasts in German, which, however, did not appear to be of systematic character[25].

The first observation to be drawn after applying both the speech type and gender factors to the corpus, was related with German read speech: both female and male subjects of highest and high aptitude displayed more variability in their realizations than the other groups. For males, such performance was closely followed by the results obtained for the otherwise recurrently aberrant behavior of the below-average speaker. The situation in English did not follow this pattern. Female native subjects replicated the performance of highest-aptitude, high-aptitude and average informants, whereby above-average, below-average and low-aptitude speakers formed the other cluster of groups producing read utterances in a conspicuously more variable way. Male realizations in English were all relatively uniform.

Most ANOVA tests, with the exceptions of the German male read-speech data ($p=.198$), produced statistically significant outcome ($p=.000$). Nevertheless no contrasts were significant in the German-language male samples ($p>.5$), and only one comparison was significant for female speakers, i.e. high-aptitude vs. average group behavior ($p=.013$), in the current condition. In English, two groups could be singled out for their atypical mean values in the *a1* parameter: low-aptitude female group productions were distinct from all the other female groups ($p<.01$); native male speakers could be characterized by a similar idiosyncratic performance ($p<.03$), which was only in statistical accord with the below-average group realizations ($p=.261$).

The final step in the examination of the PaIntE parameter *a1* consisted in analyzing spontaneous speech data split by gender. This speech type has been generally described as being less predictable and homogeneous than read speech, and it was partially supported by our previous investigation of parameter *d*. In the given data, however, i.e. *a1* values in spontaneous speech, there was a certain level of uniformity in the male realizations. The exceptions from that trend included, unsurprisingly, the below-average speaker's performance, who demonstrated less variability in German and an insufficient number of L*H tokens in English for a valid statistical

25 The peculiarities of the *a1 parameter* distribution in L*H accents in read and spontaneous speech can be seen in detail from Figure 3 and Tables 3-4 of the Appendix to the current investigation.

evaluation. Apart from that, above-average speakers produced their utterances in a less variable way than the other subject groups in English, which was especially evident in the contrast with the native speakers, who displayed the greatest degree of variability in this condition.

Female realizations showed more discrepancies between the groups. Thus, in German the otherwise comparable variability pattern was broken by the above-average and low-aptitude speakers: there was visibly less variation in the respective samples, when compared with those of the remaining groups. The English part of the corpus revealed the greatest degree of variation in the highest-aptitude group, whose performance was most closely followed by the native, high-aptitude and average groups, whereas speakers of above-average, below-average and low pronunciation aptitude produced their tokens less variably.

The cross-group contrasts were significant as rendered by the ANOVA tests ($p<.04$), but the detailed insight into the individual group oppositions only brought a few significant discrepancies, according to Scheffé post-hoc tests.[26]

Parameter *c1*

The amplitude of the rising sigmoid, i.e. parameter *c1*, quite evidently is part of the general amplitude of a given *F0* curve, which was described in the given investigation by means of the PaIntE parameter *d*. We would therefore tentatively predict somewhat similar behaviour for both of these parameters. However, only a detailed examination of the data may support or disprove this initial assumption.

Thus, the general analysis of all German data showed that highest-aptitude, high-aptitude and below-average speakers all realized their spontaneously produced and read utterances in a similarly variable way, which exceeded the degrees of variability in the remaining groups. Low-aptitude speakers, on the other hand, exhibited the smallest degree of variation in this condition. In English, the native speakers produced the most variable samples, as far as *c1* parameter values are concerned. Highest-aptitude and high-aptitude subjects approached the native-speaker performance, being, in their turn, closely followed by the below-average group, whereby the degrees of variability in the remaining groups were all lower.

The mean *F0* values lay by ~40-50 Hz in most groups in both languages, except for the above-average speakers in English (~30 Hz).

With respect to the mean values in the corpus, comparisons for both English and German were statistically significant in the ANOVA tests (p=.000). Most of the cross-group contrasts in the

26 All the cross-gender statistical contrasts in read and spontaneous speech are reflected in Tables 3-4 of the Appendix, whereby Figure 4 renders a graphical representation proper to the read speech and Figure 5 – to spontaneous samples, produced by the male and female subjects.

Scheffé tests were also significant in both languages, despite the relative equality of the mean pitch levels. More specifically, in German the low-aptitude speakers exhibited statistically significant performance in comparison with all the remaining L1 German groups (p=.000). The below-average subjects, in their turn, only agreed with the high-aptitude in terms of mean values in the *c1* parameter (p=.093); the tests with all the other groups rendered significantly different results (p<.01). In the English-language part of the corpus the native subjects displayed discordant statistical behavior as opposed to most of the non-native groups (p=.000), excluding the highest-aptitude and high-aptitude informants. Subjects of above-average pronunciation ability realized the *c1* parameter on a clearly idiosyncratically low mean level in English, which was evidenced by the statistically significant outcome, when the given sample was compared with those of all the remaining groups (p=.000)[27].

Analyzing the corpus gender-specifically allowed us to detect an overall uniformity in most of the male realizations in both languages, the only exception being the performance of the below-average male subject, whose tokens were characterized by the smallest amount of variation in both language conditions. Apart from this observation, it should be noted that the low-aptitude productions were slightly less variable in German than those of the remaining subjects. The same could in fact be found in the English realizations of the low-aptitude group as well. The above-average group sample was also less variable in English.

Female subjects produced their utterances with greater cross-group diversity. Thus, in the German data the highest-aptitude, high-aptitude and below-average group tokens were observably more variable than the remaining productions, and low-aptitude speakers departed from such performance to the greatest degree – with the most significant drop in variability. In English, the native speakers coupled with subjects of highest and high aptitude formed a cluster of groups identified by the greatest degree of variation in the parameter *c1*. The remaining female groups demonstrated less variable samples, with the lowest degrees of variability produced by the low-aptitude and above-average groups.

Almost all the ANOVA's attested to statistically significant differences in the mean values across the genders (p=.000). Most of the Scheffé tests also rendered significant inter-group contrasts in both languages under investigation (p<.05), with a few incidental exceptions[28].

After looking at the gender-specific data, we investigated the read and spontaneous speech parts of the corpus separately. The very initial examination of these samples showed that there were more discrepancies in degrees of variability across the groups in spontaneous speech, than was the case with read utterances. The most striking performance was rendered by the below-average group in spontaneous speech, whose values exposed the clearly higher degree of variability as

---

27 Refer to Figure 11 and Tables 7-8 in the Appendix for a graphical and statistical illustration of the distribution of the *c1* parameter in L*H accents in the whole experimental corpus.
28 Figure 12 and Tables 7-8 in the Appendix support the above gender-specific peculiarities in parameter *c1* of the L*H accents.

realized by these speakers in both languages at issue. Disregarding this outlying instance, we should mention that the remaining groups produced *c1* values with comparable degrees of variation in this condition. In German, only the low-aptitude and above-average informants had slightly less variability in their values than the remaining groups. English-language productions could be described as having almost equal degrees of variability in all groups, except for the low-aptitude speakers, whose sample had visibly less valuable values.

In read speech we found more homogeneity with respect to the degrees of variability across the groups. The German part of the corpus demonstrated a slightly greater degree of variation in the below-average, high-aptitude and highest-aptitude groups. The average and above-average speakers occupied a medial position in this respect, while the low-aptitude subjects' tokens were least variable in the given condition. English-language productions revealed the greatest degree of variability in the native-speaker group, whose pattern was almost replicated by the highest-aptitude informants. High-aptitude and below-average speakers produced their utterances in read speech with a marginally lower degree of variability. Finally, the most mean-centered values were to be found in the above-average and low-aptitude groups.

The inter-group contrasts were significant in the ANOVA tests ($p<.002$). In German read speech the low-aptitude group demonstrated a deviant performance differing statistically from all the remaining groups ($p<.02$). In German spontaneous speech it was also the lower-aptitude subjects who produced the aberrant mean values: the low-aptitude group disagreed with the above-average, average and below-average speakers in this respect ($p<.05$), whereas the below-average group was different from all the groups in this condition ($p=.000$). In the English read-speech data most of the contrasts were significantly different from one another ($p=.000$). Nevertheless the exception to this trend was presented by the native group as compared with the highest-aptitude and high-aptitude speakers ($p>.2$). There were fewer statistically significant contrasts in spontaneous speech, and those did not appear to carry any implications as to the cross-group peculiarities[29].

At the final stage of the investigation of *c1* parameter distribution, in the same fashion as with the other PaIntE parameters, we applied both the speech type and the gender factor to the corpus.

For read speech we found more homogeneity in the male productions in general, apart from the recurrently aberrant behavior of the below-average speaker, who produced the least variable samples of all the groups under investigation in both the investigated languages. Other than this discrepancy, almost all of the German-language samples were quite uniform. In English male read speech, the greatest degree of variability was detected in the native, highest-aptitude and high-aptitude groups, whereas the above-average speakers' productions were least variable (not taking the below-average male speaker into account).

29 The findings in read and spontaneous speech of the *c1* parameter of L*H accents are presented by Figure 13 and Tables 7-8 of the Appendix.

The female read speech data also pointed to some tendencies in the *c1* parameter distribution across the groups. Thus, in German the above-average and low-aptitude test persons displayed less variation than the remaining groups. For the English realizations, the higher-variability cluster was populated by the native-speaker samples, highest-aptitude, high-aptitude and below-average speakers. The tokens of the remaining groups were less variable in this condition.

The general ANOVA tests demonstrated statistical significance across the groups (p=.000) in read speech data. Most Scheffé tests produced results within the significance interval as well. First we analyzed female realizations in German Thus, the low-aptitude group productions in this language were only not significantly different from those of the above-average group (p=.828). For the below-average group's L1 performance in this condition it was the highest-aptitude group that produced similar mean values (p=.089).

In the male realizations in German read speech, the below-average speaker rendered a deviating sample by producing read speech utterances with a significantly differing mean *c1* parameter value level (p=.000 in all contrasts).

English data also abounded in statistically significant contrasts. In this condition, the female data indicated a significant deviation of the highest-aptitude, high-aptitude and native groups cluster from the remaining groups. More specifically, all the contrasts involving these groups produced significant results (p=.000), with the exception of the highest-aptitude vs. high-aptitude group juxtaposition (p=.535) and the native-speaker vs. highest-aptitude informants' performance (p=.401). The remaining tests yielded no significant results (p>.1). The male samples did not present any such clear pattern. Nevertheless most comparisons were also statistically significant. The below-average speaker, similarly to the read-speech productions, produced L*H targets on a statistically different mean pitch level than the other informants (p<.01), excluding the above-average group sample (p=.453).

The final part of the investigation of the *c1* parameter distribution in the corpus, i.e. in spontaneous speech, rendered perhaps the most conspicuous results in the degrees of variability across the groups. It was especially evident in female speech in both languages we examined. In German, there was a quite prominent difference in variability degrees between the high-aptitude and below-average speakers, on the one hand, who displayed the largest scattering of *c1* values, and the remaining groups, especially the low-aptitude speakers, whose tokens were most concentrated around the mean. The English-language samples retained this trend by the high-aptitude and below-average speakers, but in this condition such performance was also approached by the values produced by the native, highest-aptitude and average groups. The low-aptitude and above-average group samples were, in their turn, marked by the smallest degree of variability.

The only visible cross-group differences in the realization of the *c1* parameter in male spontaneous speech regarded the low-aptitude group and, especially, below-average subject in German, whose degree of variation was clearly lower.

The general ANOVAs attested to statistical significance in the data (p=.000). However, the Scheffé tests produced few statistically significant contrasts. Thus, no statistical trends were discovered in the German part of the corpus The English female data only contained statistically significant contrasts of the above-average group compared with highest-aptitude, high-aptitude, below-average and native speakers (p<.05). The male productions in the given language revealed significant discrepancies between the native speakers and most non-native groups, except for the highest-aptitude and below-average informants (p>.8)[30].

7.1.1.2 Simple high targets H*

The simple high tonal targets, or H* ToBI pitch accents, were one of the two most frequent categories in the experimental corpus along with the rising tones.

In general, we will follow the same procedure in the investigation of H* accents as the one adopted for L*H events. We will, however, reduce the number of PaIntE parameters for the analysis, limiting it to the description of the amplitudes of individual *F0* peaks, i.e. parameter *d*, and the peculiarities of pitch peak alignment – rendered by parameter *b*. Such a reduction is caused by the fact that no significant pitch movement is supposed to take place in an H* target. Instead, it should be realized as a plateau-like segment on a high *F0* level. Therefore, the alignment of the peak and the actual numeric representation of this pitch level (in Hertz) appear to capture the necessary information about each individual H* accent.

Parameter *d*

As mentioned in the above introduction, we will follow the same general procedure employed to the description of L*H pitch accents. Thus, we will start by investigating the peculiarities of the distribution of H* events in the whole corpus and further proceed with the analysis of the corpus split by the factors of gender and speech type, applied both separately and simultaneously.

30 Tables 7-8 of the Appendix contain the results of the Scheffé tests for read and spontaneous speech in male and female productions of L*H accents; Tables 14-15 demonstrate the variability patterns between the groups in two speech types – read and spontaneous utterances, respectively.

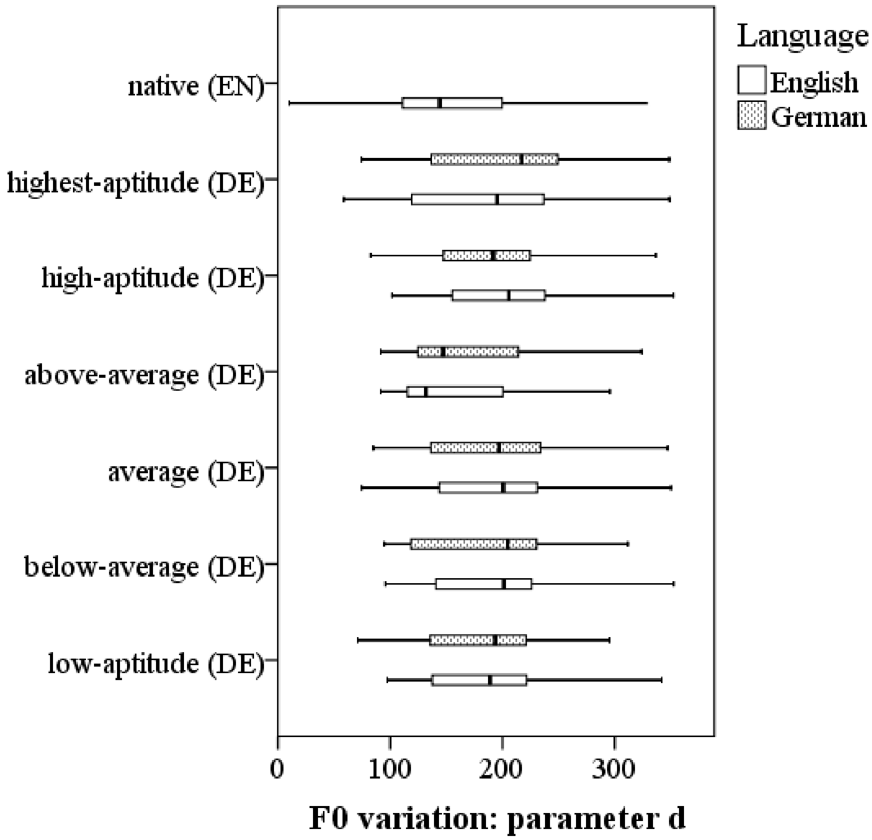


Figure 7.5: d parameter variation in H accents with regard to the F0 level in Hz in all the aptitude groups in English and German*

General corpus analysis

The initial examination of the experimental corpus revealed no striking cross-group differences with regard to the degree of variability of parameter *d*. Nevertheless some observations could be drawn for both languages at issue. In their L1 productions, most German groups exhibited comparable variation degrees, with the exception of the below-average and highest-aptitude groups, whose values were clearly more scattered.

Still another conspicuous fact for the German-language part of the corpus was related to the performance of the above-average speakers: the mean values in these realizations lay significantly lower than in all the remaining groups (~140 Hz vs. ~190-220 Hz), especially as contrasted with the highest-aptitude informants, who demonstrated the highest mean pitch level (~220 Hz).

In English, the above-average subjects behaved in the same way, i.e. realizing the H* accents on an even lower mean *F0* level than their German-language productions (~130 Hz). Interestingly, such performance was closest to the native group in this respect, as L1 English speakers also had relatively low mean pitch values, though slightly higher than those delivered by the above-average informants (~140 Hz). The other non-native groups displayed mean values at ~180-190 Hz.

As far as the degrees of variability are concerned, it should be noted that the highest-aptitude test persons had the most variable tokens of all groups, the latter sharing almost equal variation patterns.

ANOVAs yielded significant results for both languages (p=.000). The Scheffé post-hoc tests, in their turn, supported our observations on the differences in pitch levels between the groups. Significance was confirmed for the contrast between the highest-aptitude group and the other groups in German (p<.01). Above-average subjects were also identified as statistically deviant from all the other speakers (p<.01), with the exception of the low-aptitude informants (p=.114). These results were apparently triggered by the observably lower mean *F0* as realized by the above-average subjects and a clearly higher pitch level of the highest-aptitude group productions in the German-language data. The post-hoc tests carried out for English also contained evidence favoring our initial empirical observations of the mean pitch in the investigated groups.

The above-average speakers' values were therefore statistically different from the ones rendered by the remaining L1 German groups (p=.000), but not from the native-group sample (p=.819). Evidently, this fact can be explained by the greater proximity between the mean values in the native and above-average group tokens, as opposed to the higher mean *F0* in the remaining subject pool. Apart from that, the high-aptitude informants also appeared to stand out from the other groups in this condition, as all the comparisons were statistically significant (p<.03). This finding does not allow of an immediate explanation due to the fact that the mean *F0* values produced by the high-aptitude group lay on approximately the same level as those of the average, below-average and low-aptitude speakers. The highest-aptitude group sample was statistically opposed to the tokens of the high-aptitude, above-average and native subjects (p=.000): the average *F0* level of these speakers was higher than in the former two groups, but lower than the mean values of the latter informants[31].

31 Tables 11-12 in the Appendix contain all the results of the Scheffé tests discussed for the whole corpus of H* accents, as well as for all the subsequent experimental conditions: read and spontaneous speech; male and female data; and gender-specific read and spontaneous speech realizations.

Gender-specific analysis

Analyzing the data by gender allowed us to detect somewhat similar tendencies as the ones we identified during the investigation of the *d* parameter in L*H accents. Thus, in general, as with the rising tones, there was more regularity in the female part of the corpus. Another recurring pattern in female productions concerned the degrees of variability and mean pitch levels. In this respect, the average *F0* values were highest in the highest-aptitude group in both English and German and high-aptitude speakers in English (~240 Hz vs. ~210-230 Hz in the remaining groups).

As to the degrees of variability, we could state that the highest-aptitude and high-aptitude group values were marked by a greater than average degree of variability in the two languages under investigation. In German, only the average speakers produced equally variable tokens; and in the English data they approached the degree of variability in the highest-aptitude high-aptitude groups. Native informants demonstrated the same degree of variation as the two higher-aptitude groups in English. The remaining speakers realized their utterances in a visibly less variable way.

Still another observation for this experimental condition ought to be made about the mean pitch level of the L1 English female subjects: along with the average *F0* values produced by the below-average and above-average speakers, it was lower than the values in the other groups (~200-210 Hz vs. ~220 Hz, accordingly). Finally, a conspicuous fact was that the majority of the non-native female groups lowered their mean pitch by ~10 Hz when speaking English, except for the low-aptitude and highest-aptitude groups, who realized their prominent syllables on a slightly higher mean pitch in L2.

Male productions presented a more heterogeneous picture of the *d* parameter distribution. Similar to their performance with respect to the L*H accents, the high-aptitude group realized their utterances on a higher mean pitch level in both languages than the remaining subjects (~140-150 Hz vs. ~110-130 Hz, respectively).

The degrees of variability in German were comparable in most male groups, except for the below-average speaker with a much more mean-centered sample. The English part of the corpus revealed the greatest degrees of variability produced by the native and high-aptitude informants. Surprisingly, there was very little variation in the highest-aptitude group sample, similar to the tokens of above-average and below-average speakers.

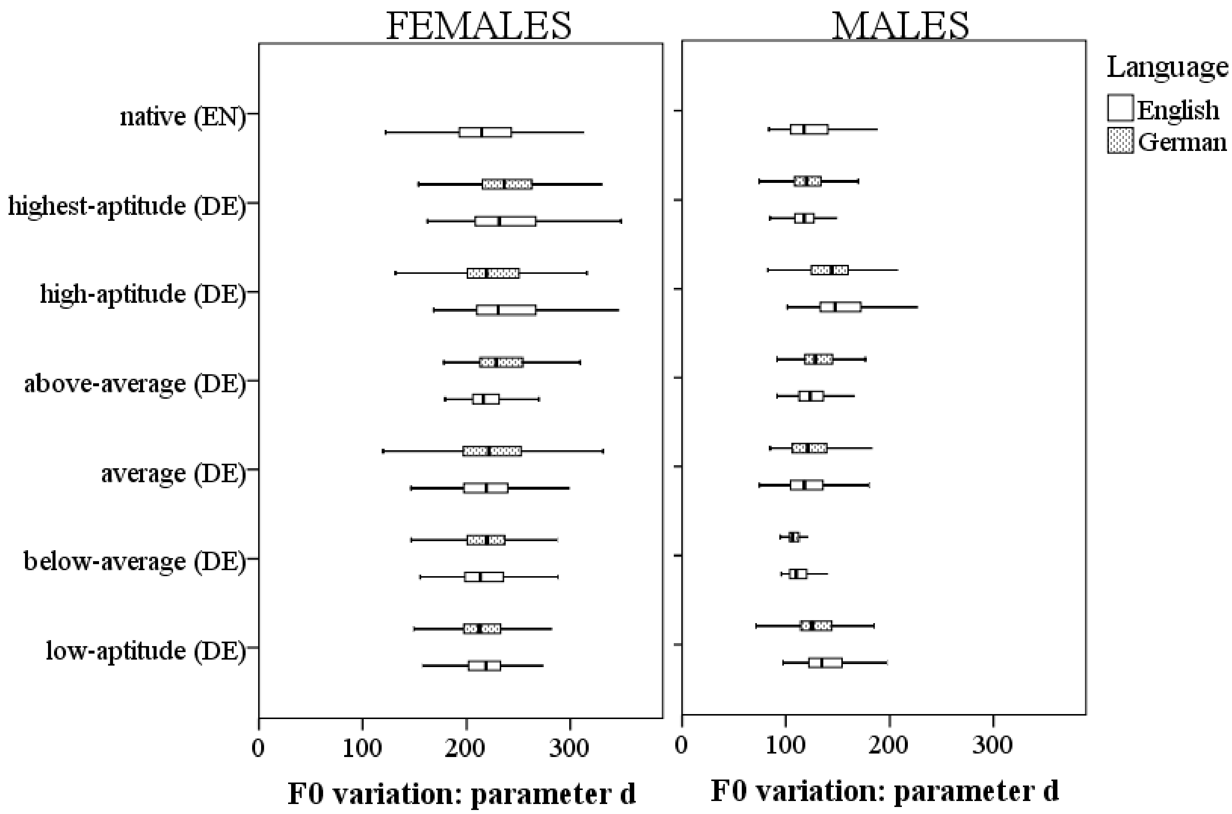


Figure 7.6: d parameter variation in H accents with regard to the F0 level in Hz in all the aptitude groups in English and German genderwise*

The general ANOVAs stated that the differences between the groups were significant (p=.000). A more detailed insight into the data was provided by the Scheffé tests, which contained some evidence for the observations presented above. More precisely, in the German female productions two groups could be singled out as deviating statistically from the rest of the subject pool: highest-aptitude female speakers differed in their mean pitch level from all the groups (p=.000), excluding the above-average subjects (p=.528); and informants of low pronunciation aptitude displayed values departing from those of the highest-aptitude, high-aptitude, above-average and average test persons (p<.01).

The clearly higher average *F0* of the male high-aptitude informants in their L1 speech made their sample statistically different in all contrasts ($p<.02$); the same effect was caused by the below-average speaker's lower than average pitch level ($p=.000$).

The English part of the experimental corpus only made the statistical distinction of the high-aptitude female group against highest-aptitude, above-average, average and below-average group productions evident ($p<.03$). In the male realizations the high-aptitude speakers once again presented a statistically outstanding sample due to their higher mean *F0* ($p<.01$), which was only comparable to the low-aptitude group tokens ($p=.882$). The latter speakers, in their turn, also differed from the rest of the subject pool, not including the high-aptitude informants, in their average pitch level ($p<.05$).

Figure 7.6 as well Tables 11-12 of the Appendix are to be referred to for a graphical and statistical evidence of the empirical observations on the gender-specific distribution of *d* parameter values in H* accents.

Corpus analysis by speech type

Separate analysis of read and spontaneous speech yielded still another perspective onto the corpus. Predictably, read samples were characterized by greater homogeneity than the spontaneously produced utterances with regard to the degrees of variability.

More specifically, the German-language realizations all exhibited comparable variation patterns across the groups, the below-average and highest-aptitude samples being marked by a somewhat greater degree of variability. In English only the highest-aptitude group presented a sample more variable than those of the other informants.

The mean *F0* values also lay within the same scope in both languages at issue (~200 Hz), but the native-speakers' utterances and those produced by the above-average group represented an exception to this general trend (~130 Hz for English and ~150 Hz for German). This appears to be a recurring phenomenon due to the earlier observation based on the whole corpus, as well as the gender-specific analysis demonstrating generally lower pitch values in these two groups.

More discrepancies between the groups were discovered in spontaneous speech, both with respect to the degrees of variability and the mean *F0* levels.

Regarding the first issue, we found that in German the high-aptitude and below-average group tokens were less variable than the respective data produced by the other groups, with the highest-aptitude informants realizing their utterances in the most variable way. English-language sample values of the above-average group were distinguished by the greatest degree of concentration around the mean.

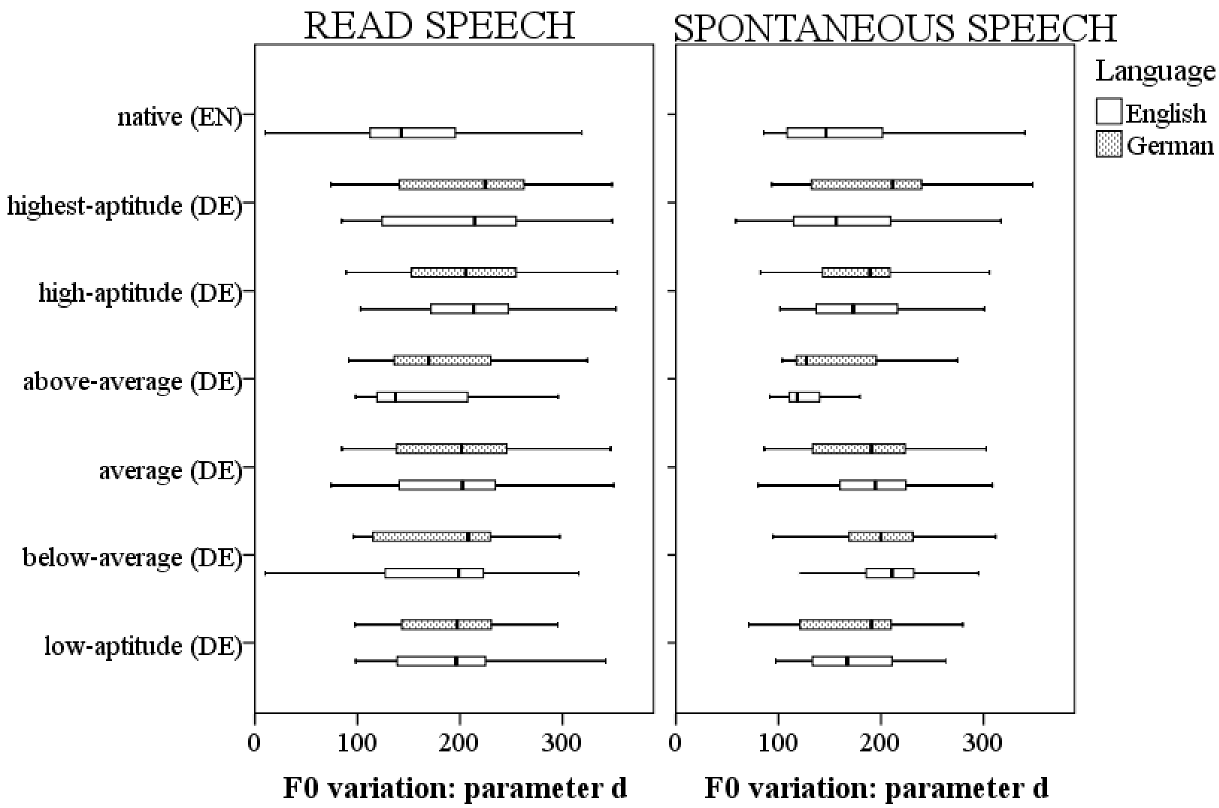


Figure 7.7: d parameter variation in H accents with regard to the F0 level in Hz in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

Below-average and average speakers followed with gradually increasing degrees of variation in the *d* parameter tokens. Low-aptitude and, unexpectedly, high-aptitude subjects produced equally variable samples. And, finally, the highest-aptitude and native-speaker groups had the most variable values in their productions.

Observations of the mean *F0* level in spontaneous speech were also quite conspicuous. The highest values hereby were found in the below-average group (~210 Hz). Somewhat lower was the average pitch of the average subjects (~190 Hz). High-aptitude and low-aptitude speakers shared the same mean token values (~170 Hz).

The highest-aptitude group most closely approached native-like performance in this condition (~140 vs. ~150 Hz). Lastly, the lowest *F0* values were found in the above-average group (~120 Hz).

ANOVA tests confirmed that the differences between the subjects were statistically significant (p=.000). Scheffé post-hoc comparisons of the German data singled out the highest-aptitude and high-aptitude groups in read speech as the speakers differing from the other informants to the greatest extent (p<.05). The only non-significant contrast for the highest-aptitude subjects was that against the high-aptitude informants (p=.737). For the latter an additional significant comparison was discovered when opposed to the mean values of the average test persons (p=.253). The remaining contrasts did not deliver any significant results. German spontaneous speech presented a less clear pattern as to the mean *F0* values in statistical sense. However, the lowest mean *F0* in the above-average group was manifest through the significance of p-values in Scheffé tests in all comparisons (p<.02).

The Scheffé results for English singled out the highest-aptitude, above-average and native groups as the speakers deviating consistently from the other samples (p<.02) in read speech.

At the final stage of *d* parameter analysis in H* accents we analyzed the data by both speech type and gender.

Corpus analysis by speech type and gender

Gender-specific read-speech materials could be described as somewhat mimicking the general tendencies discovered for the male and female productions. This fact was most strongly confirmed by the female realizations. Thus, investigating this segment of the corpus in read speech, we could reiterate our earlier observation that the samples produced by the highest-aptitude and high-aptitude female groups were marked by the greatest degrees of variability in both languages in question. This pattern was replicated by the average group in German and the native speakers in English. Less variability was detected in the remaining samples. We could also state the recurrence of the mean *F0* level pattern across the groups. In this respect, highest-aptitude and high-aptitude female subjects could also be identified by their higher mean pitch level than the other subjects in both English and German. The latter language was characterized by higher mean pitch in this condition, and the difference between the two higher-aptitude groups and the remaining test persons was not so pronounced (~250 Hz vs. ~230-240 Hz, respectively). The mismatch in English was slightly greater (~240 vs. ~210-220 Hz). Hereby, the native, above-average and below-average informants could be grouped into a cluster of speakers with the lowest mean pitch in English (~210 Hz) – a tendency spotted previously for the whole experimental corpus split by gender. One final note ought to be made regarding the general lowering of the mean pitch by most non-native subjects when producing English utterances.

Male realizations of read-speech utterances were reminiscent of the general trends as well. The high-aptitude speakers were once again distinct due to their highest mean pitch in both English and German (~140-150 vs. ~100-130 Hz). In German, they were only approached by the low-aptitude and above-average groups (~130 Hz). Looking at the degrees of variability in male samples, we could see that in German there was almost no difference between most of the groups in this sense – a trend not supported by the idiosyncratic performance of the below-average subject, whose values were extremely concentrated around the statistical mean. English productions of the native and highest-aptitude groups had the greatest degree of variability. The rest of the subjects exhibited less variation in the *F0* peak amplitudes in English.

The above empirical observations found some support in the corresponding statistical tests. ANOVA results were all statistically significant (p=.000). Furthermore, the visibly higher average *F0* peak values produced by the highest-aptitude and high-aptitude female speakers were found as being statistically different from the values pertaining to the remaining informants in German-language data (p=.000), except for the above-average group (p>.1).

In the male German-language part of the corpus in read speech, we could once again confirm the idiosyncrasy of the below-average subject in statistical terms, as all the comparisons rendered p-values below the significance level (p<.01).

The English data also separated the highest-aptitude and high-aptitude female informants, who had the highest mean *F0* values, from the rest of the subjects: all contrasts involving these groups were significant (p=.000), but for their comparison with one another (p=.889). All the other tests yielded no significant results for female speakers in English (p>.3).

Clearly deviating mean pitch levels of the male subjects of high and low pronunciation aptitude in English caused statistically significant p-values when these samples were opposed to the remainder of the subjects (p=.000). The only other significant comparison in this condition involved the low-aptitude and below-average groups (p=.029).

The investigation of spontaneous speech by gender finalized the exploration of the *d* parameter in H* accents. The general gender-proper tendencies were, but less prominently, present in this part of the corpus as well. German female speech thus revealed a slightly different variation pattern, as compared to the observations made for the whole corpus and read utterances. Average and below-average group productions contained the greatest degrees of variability here – a pattern most closely followed by the highest-aptitude group, whereas the other samples were less variable in the amplitudes of the *F0* peaks. In English, only the highest-aptitude and average female speakers had slightly more variation in the tokens, and the above-average test person demonstrated, on the contrary, a little less variability than all the other groups. There were also no striking contrasts in female speech with regard to the average pitch levels. The few exceptions from this general observation were related to the highest-aptitude speakers' behavior in German, which involved producing marginally higher mean *F0* values than the other groups (~220 vs. ~200-210 Hz). High-aptitude and low-aptitude informants, on the other hand, realized their spontaneously produced utterances on the lowest pitch level of all groups in German (~200 Hz). In English, all mean values fluctuated between around 200 and 210 Hertz.

Investigating male spontaneous speech, we found almost equal degrees of variability and similar mean pitch levels in most groups, but with a few mismatches. Those included, first of all, the strikingly smaller degree of variability and a slightly lower pitch (~100 vs. ~110-130 Hz) in the realizations of the below-average speaker. Secondly, the high-aptitude male group productions once again revealed a tendency for the relative maximum in this condition in both English and German realizations (~130 Hz), replicating the finding for the whole corpus and read speech in this group. The only other group delivering such relatively high mean *F0* values were the low-aptitude test persons in English. For the degrees of variability, the only prominent observation was a visibly greater degree of variation in the German productions of the high-aptitude speakers. The other samples were almost indistinguishable from one another in this sense.

Cross-group comparisons were significant according to the general ANOVA (p=.000). The Scheffé post-hoc tests, in their turn, provided some solid ground to our earlier observations. Thus, the generally higher mean *F0* realized by the highest-aptitude female informants in German was described as being statistically different from the pitch levels produced by most other female groups (p=.000) with the exception of the above-average speakers (p=.528). The male realization in the German-language condition supported the empirical observation about the high-aptitude speakers' aberrant behavior as compared with all the other groups (p<.03). The below-average speaker produced values differing significantly from all the subjects, excluding the average and low-aptitude speakers (p>.1).

Most contrasts in English female spontaneous speech did not yield any significant results. Male speakers produced relatively indistinguishable mean *F0* values in English, with the one exception being the performance of the high-aptitude and low-aptitude groups, whose average pitch level was visibly higher. And this fact was confirmed by the corresponding Scheffé procedure (p<.01 in all contrasts, except for the juxtaposition of these two groups in question).

Figure 7.8 and Tables 11-12 of the Appendix contain graphical and statistical evidence for the gender-specific description of *d* parameter value distribution in H* accents in read and spontaneous speech.

Parameter *b*

The final parameter under investigation with respect to the H* accents was the alignment of the *F0* peak within a three-syllable frame, covering the accented syllable and its immediate left and right context.

General corpus analysis

The initial observation coincided with the one we made for L*H accents: in general, there was more variation in the *b* parameter in the German data than it was observed for English. The inter-group comparison in this regard in German revealed a slightly greater degree of variability as realized by the speakers of highest, above-average and below-average aptitude. In English, all groups were almost indistinguishable from one another in their degrees of variation, with just a slightly increased variability level produced by the highest-aptitude test persons.[32]

Gender-specific analysis

In the cross-gender investigation of the alignment parameter, we could reiterate the trend for a greater variability in the German-language samples both in the female and male realizations. Looking at the female data separately, we discovered an almost uniform variation pattern in the English-language tokens.

The German productions, on the other hand, singled out the below-average, as well as the highest-aptitude female speakers, as the two groups exhibiting higher than average variability degrees.

An unexpected finding was related to the male speech in English. In this condition, the lowest degree of variability was discovered within the lowest-aptitude informants, who were followed by the below-average and native groups with slightly more variable samples in this condition. The remaining groups realized English utterances with almost equal degrees of variability. A similar pattern was found in the German realizations, where the low-aptitude and below-average group samples were marked by a decreased degree of variation as compared to the rest of the subject pool[33].

32 Figure 16, Appendix illustrates the above findings in parameter *b* of H* accents in the whole corpus.
33 Figure 17, Appendix, illustrates gender-specific peculiarities in the realization of the *b* parameter in H* accents.

Corpus analysis by speech type

We could also draw some speech-type specific trends during the investigation of read and spontaneously produced utterances. Thus, it was found that English read speech was fairly homogeneous as to the degrees of variability of parameter *b* in H* accents, produced by subjects in all the six aptitude groups at issue.

Only the above-average speakers had slightly more variable H* accent tokens. In the German part of the corpus, the highest-aptitude, above-average and below-average groups were marked as distinct, due to their greater degree of variability in the *b* parameter than the remaining subjects.

German-language spontaneous utterances, on the contrary, did not contain many cross-group mismatches in degrees of variability. However, the samples of the highest-aptitude and high-aptitude groups were slightly more variable. For English spontaneous speech, we could point out two groups as falling out of the otherwise uniform variability pattern, i.e. low-aptitude and above-average subjects, who exhibited values visibly more concentrated around the mean than the remaining informants[34].

Corpus analysis by speech type and gender

Finally, we analyzed the data taking into account both the speech type and gender factors. For the female speakers in read speech, we could only spot one instance of idiosyncratic performance. More specifically, the below-average subjects had visibly more variability in their values than all the other groups in German. Apart from that, the samples had comparable variability degrees in both languages.

There was less agreement in the male data in read speech. Indeed, both the English and German realizations presented a heterogeneous variation picture. In the former language, above-average and average test persons produced more scattered values than the other groups, which was especially true for the low-aptitude subjects with the lowest degree of variability. These speakers also had the least variable sample in German, whereby below-average, average and high-aptitude speakers had an equally higher degree of variation in this condition. Lastly, the tokens of the highest-aptitude and above-average groups demonstrated the greatest dispersion of values around the mean.[35]

Spontaneous speech also revealed more consistency and homogeneity in female realizations. Here, the German tokens in female performance were identified by an almost equal degree of variation by all speakers with the exception of the highest-aptitude group and the above-average female test person, whose samples were visibly more variable. In English, degrees of variability

---

34 Figure 18, Appendix, refers to *b* parameter variation by speech type in H* accents.
35 Figure 19, Appendix, reflects gender-specific findings of parameter *b* in read speech in H* accents.

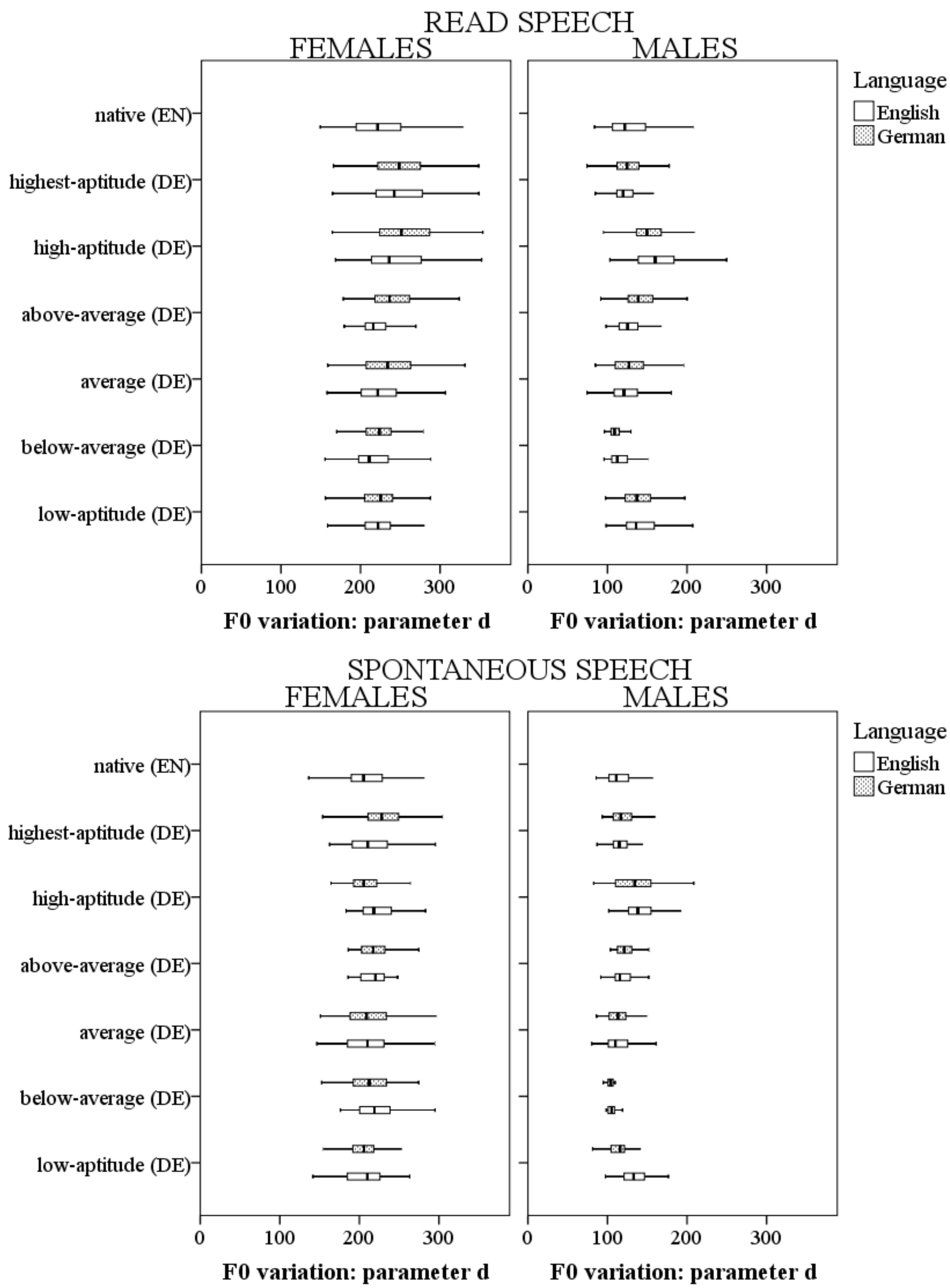


Figure 7.8: d parameter variation in H accents with regard to the F0 level in Hz in all the aptitude groups in English and German read and spontaneous speech by gender – male versus female realizations*

across the groups were also comparable. In this condition, however, average speakers were distinguished from the remaining subjects by their greater degree of variation in the *b* parameter in H* accents.

When we looked at the male data, we could not identify any clear-cut pattern with respect to variability. The German part of the corpus highlighted the average speakers as the most variable sample, which was then followed by the highest-aptitude and high-aptitude groups with a slightly decreased degree of variability. The remaining three German groups were marked by the smallest degree of variation in German spontaneous speech. Yet another pattern was proper to the English-language utterances was as follows: the below-average speaker had the most variable set of *b* parameter tokens. Least variable, on the other hand, were the samples produced by the speakers of low and above-average aptitude, as well as native speakers. The other speakers had equal levels of variability in their values.[36]

Most statistical ANOVA tests pertaining to the distribution of parameter *b* values of the H* accents yielded non-significant results ($p>.05$). The only exceptions to this general trend were the following conditions: all English data without regard to speech type and gender; cross-gender comparison – there was significance in both female and male samples; in English read speech and in read speech tokens compared gender-specifically ($p<.05$). Accordingly, most post-hoc Scheffé comparisons were also non-significant, with only sporadic instances of significance in the English data.[37]

7.1.1.3 Simple low targets L*, falling tones H*L and falling-rising tones L*HL

The number of tokens in L*, H*L and L*HL did not suffice for a valid statistical examination with the application of all factors, we limited the current investigation to the analysis of the whole corpus in general, as well as female and male realizations taken separately.

As was the case with the most frequent accents, the PaIntE parameters for a detailed investigation at this stage were selected based on their pertinence to the description of a given category. Thus, all events were treated in terms of the amplitude of the *F0* curve, or parameter *d*, and the alignment of the pitch peak within the accented syllable and in its immediate proximity – parameter *b.* In addition to that, we looked at the peculiarities of the production of falling sigmoids in H*L and L*HL accents, rendered through the corresponding values of parameters *a2* (the velocity of the fall) and *c2* (the amplitude of the fall).

[36] Figure 20 of the Appendix reflects the gender-specific variation peculiarities of parameter *b* in spontaneous speech in H* accents.
[37] Tables 9-10 of the Appendix present a detailed overview of Scheffé test results for parameter *b* of H* accents in all experimental conditions.

Parameter *d*

General corpus analysis

The initial look at the distribution of the *d* parameter values in the corpus allowed us to conclude that there was generally more regularity with respect to the degrees of variability in the German data than in English.

Rising-falling tones L*HL

More specifically, the rising-falling tones were marked by almost equally variable samples in all groups in German, with the exception of the low-aptitude speakers who produced L*HL accents with a clearly smaller degree of variation. As far as the pitch levels are concerned, it should be noted that it was highest in the below-average group (~220 Hz). Highest-aptitude and average speakers exhibited a slightly lower *F0* of the rising-falling targets (~210 Hz). The mean *F0* values in the other groups lay somewhat lower at ~180-200 Hz.

Notably, this trend was to a certain extent carried over into the English-language realizations. Only here the distinction of the above-average and low-aptitude informants as to their mean *F0* values was even more pronounced: ~170 and ~190 Hz, accordingly, in those groups against ~220-290 Hz in the remaining subject pool. The closest pitch level to these two outlying groups was produced by the below-average subjects (~220 Hz), whereby the remaining L1 German test persons shared an almost equal mean *F0* level of ~240-250 Hz. Still another surprising outlier was represented by the native-speaker group sample with average L*HL values exceeding the highest level in the L2 speakers' realizations by ~40 Hz, i.e. the mean *F0* was registered at ~290 Hz in this group.

With regard to the degrees of variation in L*HL accents in English, it should be pointed out that the low-aptitude group presented the least variable sample, rather mimicking their performance in German. At the other end of the variability spectrum, the highest-aptitude speakers demonstrated the greatest spreading of values around the mean. Slightly lower, but to an equal proportion, was the degree of variability in the average and below-average groups. Producing values still a little more concentrated around the mean, above-average, high-aptitude and native speakers formed the final cluster in this condition.

The ANOVA tests supported the empirical observations about the distribution of the *d* parameter values in the L*HL accents in a way that statistical significance was postulated for the comparisons in both English and German (p =.000). The post-hoc Scheffé tests established a statistical distance of the low-aptitude group from the majority of speakers in German (p<.03), excluding the above-average speakers who produced values within the same average scope (p=.994). The rest of the cross-group comparisons were not significant (p>.1).

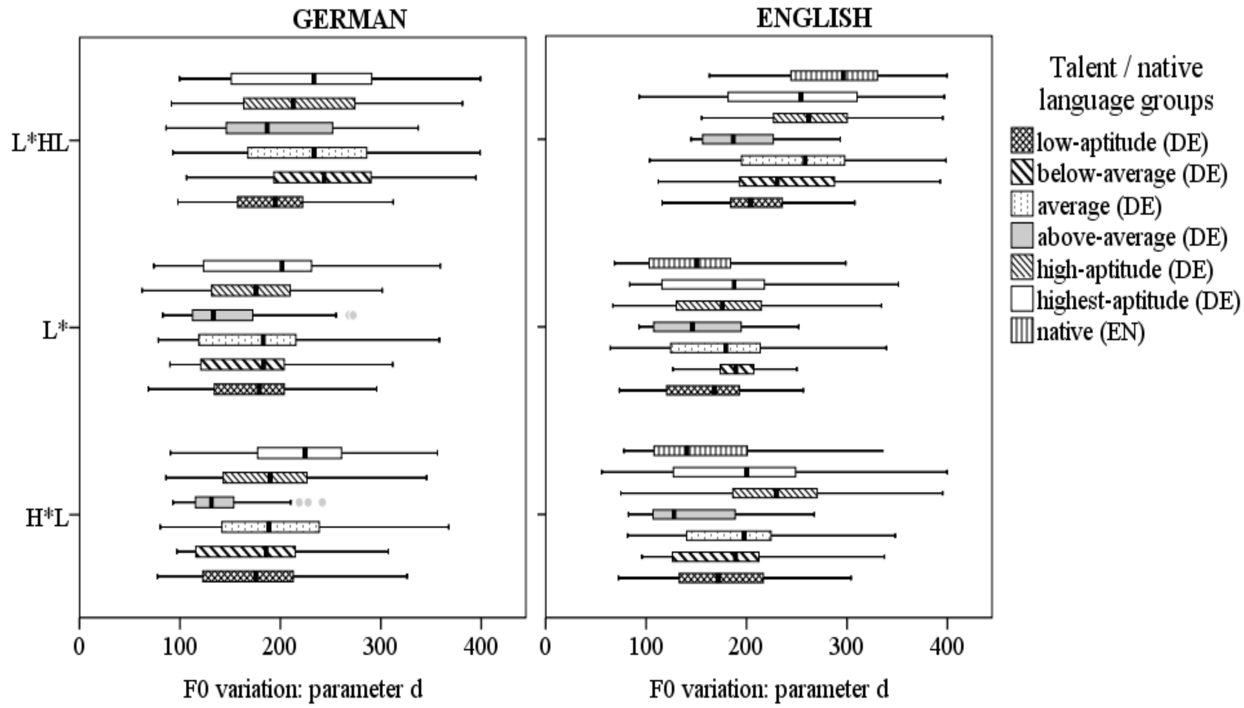


Figure 7.9: d parameter variation in less frequent accents (L, H*L, L*HL) with regard to the F0 level in Hz in all the aptitude groups in English and German*

In the English data, the low-aptitude group sample was statistically different from the values produced by all the other groups ($p<.02$).

Simple low targets L*

We next investigated the simple low tonal targets L* in the entire corpus. This analysis yielded fewer distinctions between the groups in the degrees of variability. There were only a few cases of departure from the general frequency trend. In German, for instance, the above-average group presented the least variable sample, whereby the degrees of variability were relatively comparable in the remaining groups. Nevertheless it should be noted that of these groups the highest-aptitude speakers were still marked by the greatest degree of variation, which was most prominent when contrasted with the low-aptitude subjects' sample – with the greatest concentration of values around the mean of most groups, excluding the above-average test persons.

As for the mean *F0* level, it should be pointed out that similar values were shared by most groups (~180 Hz). Outlying mean pitch levels were found in the highest-aptitude (~200 Hz) and above-average groups (~130 Hz).

The latter speakers transferred this pattern into English, by producing lower mean values than the remaining L1 German groups (~140 vs. ~170-190 Hz). Unexpectedly, the native speakers also realized their L* accents on a similar lower mean pitch level. The variability degrees in this condition were almost equal in all groups, but for the subjects of below-average aptitude, whose sample was clearly less variable than the rest.

Inter-group differences were significant according to the respective ANOVAs (p=.000). As already mentioned in the empirical description, the highest-aptitude and the above-average speakers produced outlying mean values in German, which supported by the Scheffé findings. The latter group differed significantly from the remaining subject pool in statistical sense (p<.01). The highest-aptitude speakers, in their turn, also delivered a distinct sample, which was only statistically comparable with the values produced by the below-average group (p =.074). In the English data, only the L1 English informants appeared to consistently deviate from the general statistical trend: their values were only indistinguishable from the ones produced by the above-average group (p =1), as well as the low-aptitude subjects (p =.676).

Falling tones H*L

We finally investigated the falling H*L accents in the whole corpus. In general, we found more consistency in the variability degrees than in the average pitch levels. Similarly to the L* accents, native and above-average speakers shared an almost equal lowest *F0* level of all groups in the falling tones (~130-140 Hz). Somewhat higher were the pitch values produced by the low-aptitude, below-average, average and highest-aptitude groups (~170-195 Hz). Finally, the high-aptitude group appeared to be the outstanding sample with the mean *F0* level of around 220 Hz.

The German data revealed comparable degrees of variability and mean pitch levels at ~180-190 Hz in most groups. The two exceptions to the given pattern in this condition included the higher *F0* level within the highest-aptitude informants (~220 Hz), as well as a visibly smaller degree of variability and a lower pitch of the above-average informants (~130 Hz).

The ANOVA tests seconded the fact that the inter-group differences were significant (p =.000). According to the Scheffé comparisons, in German the abnormally higher pitch level of the highest-aptitude group and the visibly lower mean *F0* values produced by the above-average speakers resulted in statistical separation of these groups as contrasted to the remaining subjects (p<.02). The only other statistically significant contrast was between the high-aptitude and above-average groups (p=.000). The English data also singled out above-average group statistically, due to their low mean *F0* (p =.785), distinct from the remaining groups' values

($p<.03$). Also outstanding were the mean values of the high-aptitude speakers with their clearly higher than average *F0* level ($p=.000$), as compared to all the other English-language samples. The highest-aptitude subjects only coincided statistically with the average and below-average groups ($p>.1$).

Figure 7.9 presents the findings in the distribution of L*, L*HL and L* accents; Tables 13-14 of the Appendix render statistical cross-group information produced by the Scheffé tests.

Gender-specific analysis

At the final stage of the investigation of *d parameter* values in less frequent accents, we compared the gender-specific samples. As a result, we could state, in general, that the realizations exhibited a more heterogeneous picture than before splitting the corpus into male and female speech, both in the English- and German-language productions.

Rising-falling tones L*HL

The female realizations of the rising-falling tones L*HL in German were fairly uniform with respect to the degrees of variability. Here, the highest-aptitude, high-aptitude and average speakers exhibited almost equally variable samples. The two groups differing from this trend were the below-average speakers with the greater and above-average and low-aptitude subjects – with a lower degree of variation. The mean *F0* values lay on almost the same level in all groups (~250 Hz). The only outstanding mean pitch values were produced by the low-aptitude subjects (~200 Hz).

The English-language syllables carrying the rising-falling targets, produced by female subjects, appeared to suggest a pattern of gradually ascending mean *F0* levels with increasing degrees of language aptitude. Thus, the lowest mean pitch was found in the low-aptitude group tokens (~200 Hz), and the highest – within the L1 English speakers (~280 Hz). There was, however, one group sample departing from this pattern. More specifically, the above-average informants realized their L*HL accents on a lower mean level than the average group. Apart from that, the above-average female subjects did not fit into the general variability pattern in this condition: the values in that sample were clearly more concentrated around the mean than in the remaining groups. Regardless of this irregularity, the other samples were almost equally variable, except for the below-average group utterances, which had visibly more variation in the *d* parameter values of L*HL accents.

Statistically, the differences between the female groups were significant according to the ANOVA results in both languages under investigation ($p=.000$). The post-hoc tests, in their turn, supported most of the above empirical observations about cross-group peculiarities. More specifically, the above-mentioned observably lower mean pitch level in the low-aptitude group in German was found to be statistically different from the remaining samples, as stated by the Scheffé tests (*p*=.000). The statistical picture in English was less straightforward. Here, the lower *F0* level of the low-aptitude subjects also resulted in significant comparisons with all the groups (*p*<.04), with the exception of the above-average speakers (*p*=.614). A second trend involved the below-average subjects, who also appeared to deviate statistically from all groups (*p*<.04), but for the above-average and average subjects (*p*>.2).

When we turned our attention to the male data, the German-language productions appeared to contain more regularity, especially with regard to the variability degrees: they were relatively equal in all groups, with just the average speakers deviating visibly with their more variable sample. The same mean *F0* level was also shared by most of the groups (~150 Hz). The highest-aptitude and below-average speakers, however, demonstrated lower mean pitch (~120 and ~100 Hz, accordingly).

Several cross-group differences were discovered in the male samples in the English-language utterances. In this condition, the average group's values were scattered around the mean to the greatest extent, which was followed by the performance of subjects of highest pronunciation aptitude. The native, above-average and low-aptitude speakers produced similarly less variable tokens. Finally, the remaining two groups had the smallest amount of variation in their values.

High-aptitude and below-average subjects demonstrated idiosyncratic performance in the mean *F0*: the former clearly exceeded the levels produced by the other groups (~220 Hz), whereas the latter, vice versa, realized their utterances on an observably lower pitch (~120 Hz). The remaining mean *F0* values fluctuated between ~150 and ~180 Hz.

The ANOVA tests were significant for both languages in this condition (*p*<.01), although the significant concrete cross-group contrasts reflected by the Scheffé tests were not numerous.

For instance, in German no cross-group contrast was significant in male speech (*p*>.1). In the English data, only the high-aptitude group appeared to consistently deviate from the other samples (*p*<.03), except for the contrast with the average as well as native speakers (*p*>.3).

Simple low targets L*

When we proceeded with the analysis of simple low targets L*, starting with the female German-language part of the corpus, we discovered another instance of idiosyncrasy exhibited by the above-average speaker – a visibly higher degree of variability than in the other groups. Pitch-level differences in this condition were almost indistinguishable: with the general mean of ~200 Hz and a slight increase of that value by the high-aptitude and highest-aptitude speakers (~210 Hz). In English, we found relative agreement with respect to the degrees of variability in all the groups, with the exception of the below-average and especially above-average test persons with less variable samples.

The mean *F0* levels in L* targets in English fluctuated between ~180 and ~210 Hz, with the maximum values produced by the speakers of highest and high pronunciation aptitude.

Most of the above observations about female productions were supported by the statistical significance of ANOVA cross-group tests in English and German (p=.000). Scheffé tests yielded predominantly significant results as well, as far as inter-group comparisons are concerned (p<.05). The non-significant exceptions did not appear to be systematic.

L* accents in male realization did not suggest any clear pattern either in English or German. The latter revealed an equal mean *F0* level of ~110 Hz by most speakers, excluding the high-aptitude and above-average informants (~130 Hz). These two groups also produced the most variable samples, along with the low-aptitude test persons. The lowest degree of variability was delivered by the highest-aptitude and below-average speakers.

Two main observations could be drawn from the male data in the English language, both of them involving the high-aptitude group, who, first of all, realized the L* targets on a generally higher pitch (~140 Hz vs. ~120 Hz). Secondly, the degree of variability in these tokens was much higher than in the other groups.

In general, the above-described differences in male data appeared to be significant in statistical terms (ANOVA, p=.000), whereby the Scheffé post-hoc corrections provided more detailed ground. Thus, the only sample consistently deviating from those rendered by most other groups pertained to the high-aptitude informants: contrasts against most other groups in English and German were significant (p<.03).

Falling tones H*L

Further on, we analyzed the gender-specific peculiarities of *d* parameter distribution in H*L accents. The very first inspection of the tokens in this condition indicated a certain tendency in female realizations in English. More specifically, as mentioned for several previous conditions, the samples produced by the highest-aptitude and high-aptitude informants could be grouped together with the native-speaker utterances with respect to the degree of variability, which was clearly higher in these groups. However, the three samples in question reflected different mean *F0* levels: the high-aptitude group thus produced the highest values (~ 250 Hz), whereby the corresponding levels in the highest-aptitude and native groups were visibly lower (~ 220 Hz and 210 Hz, respectively). Unexpectedly, the only other group matching the female native speakers in this respect were the low-aptitude informants – with almost equal average pitch values. The remaining test persons exhibited lower mean *F0* (~ 200 Hz) as well as smaller degrees of variation in English.

Several observations could be drawn from the female productions in the German language. First of all, evident was the higher degree of variability in the highest-aptitude, high-aptitude and, especially, average-group utterances. The remaining samples were less variable, which was most conspicuous of the above-average female subject. The mean *F0* level was almost equal in the majority of groups (~210 Hz), with the exception of highest-aptitude speakers, who produced their tokens on a higher pitch (~230 Hz). The given pattern therefore disagreed with the one described for the English language above.

ANOVAs suggested that the differences between the female groups were significant in English as well as German (p=.000). However, the Scheffé test only put forward the juxtaposition of the highest-aptitude versus the average, below-average and low-aptitude groups as significant in German (p<.01), most probably due to the higher-than-average mean *F0* produced by this group. In English, the sample produced by the high-aptitude speakers was found to be statistically different from most of the remaining groups (p<.02), except for the highest-aptitude subjects also producing relatively high mean *F0* values (p=.132).

Finally, we investigated the distribution of the *d* parameter in H*L accents as realized in the male samples. The most conspicuous fact was discovered in English male speech, where the high-aptitude subjects exhibited an abnormally high degree of variability, as well as a notably higher mean *F0* level (~160 Hz vs. ~130 Hz in the low-aptitude group and ~110 Hz in the remaining groups).

No striking contrasts with respect to the degrees of variability could be found in male speech in German. Nevertheless a visibly lower degree of variability was discovered the below-average test person as opposed to all the other groups. As far as the mean *F0* is concerned, it should be noted that the highest level was registered for the high-aptitude informants (~130 Hz vs. ~110 Hz in the remaining groups) – the fact that has been mentioned for this group previously.

The above differences in mean *F0* levels were significant according to the ANOVA tests (p=.000). The Scheffé attested to statistical similarity between the mean values in most contrasts in German. In the English samples, the clearly higher mean *F0* demonstrated by the high-aptitude group resulted in the statistical significance of all contrasts involving these subjects (p=.000)[38].

38 The cross-gender findings in less frequent accents – L*, L*HL, H*L – are summarized graphically in Figure 7.10 and statistically – in Tables 13-14 of the Appendix.

Parameter *b*

L*, L*HL and H*L accents were further analyzed with respect to the alignment of *F0* peaks in them – parameter *b*.

General corpus analysis

Rising-falling tones L*HL

The very first look at the distribution of this PaIntE parameter presented a rather uniform pattern in the rising-falling L*HL tones both in English and German. Not only were the degrees of variability similar, the mean values were also almost equal. Thus, the speakers appeared to realize the peaks in these accents around the first post-accented syllable. Also notable was the fact that there was visibly less variability in English data as opposed to the German-language realizations.

Accordingly, the ANOVA results supported this homogeneity in the *b* parameter values by postulating no statistical significance between the samples ($p>.1$).

Simple low targets L*

Relatively uniform was also the variability pattern in L* accents, especially in German, where the degrees of variability between the groups were almost equal; so were the mean values in the aptitude groups. In the English realizations the degrees of variation were also quite similar. The above-average group, however, departed from this general trend by exhibiting a conspicuously more variable sample in this respect, coupled with a slightly lower mean level. Notably, the highest-aptitude group also realized L* targets with lower mean values.

Concurring with the above empirical observations, the inter-group differences in English were found to be significant according to the ANOVA tests ($p=.001$). The corresponding comparison of the German data was not significant ($p=.512$). The post-hoc Scheffé tests did not reveal any significant contrasts in both languages whatsoever ($p>.1$).

Falling tones H*L

The investigation of H*L accents with relation to *F0* peak alignment was not as uniform as in L* accents. In fact, it yielded a few notable mismatches between the samples, both with respect to the degrees of variability and mean levels.

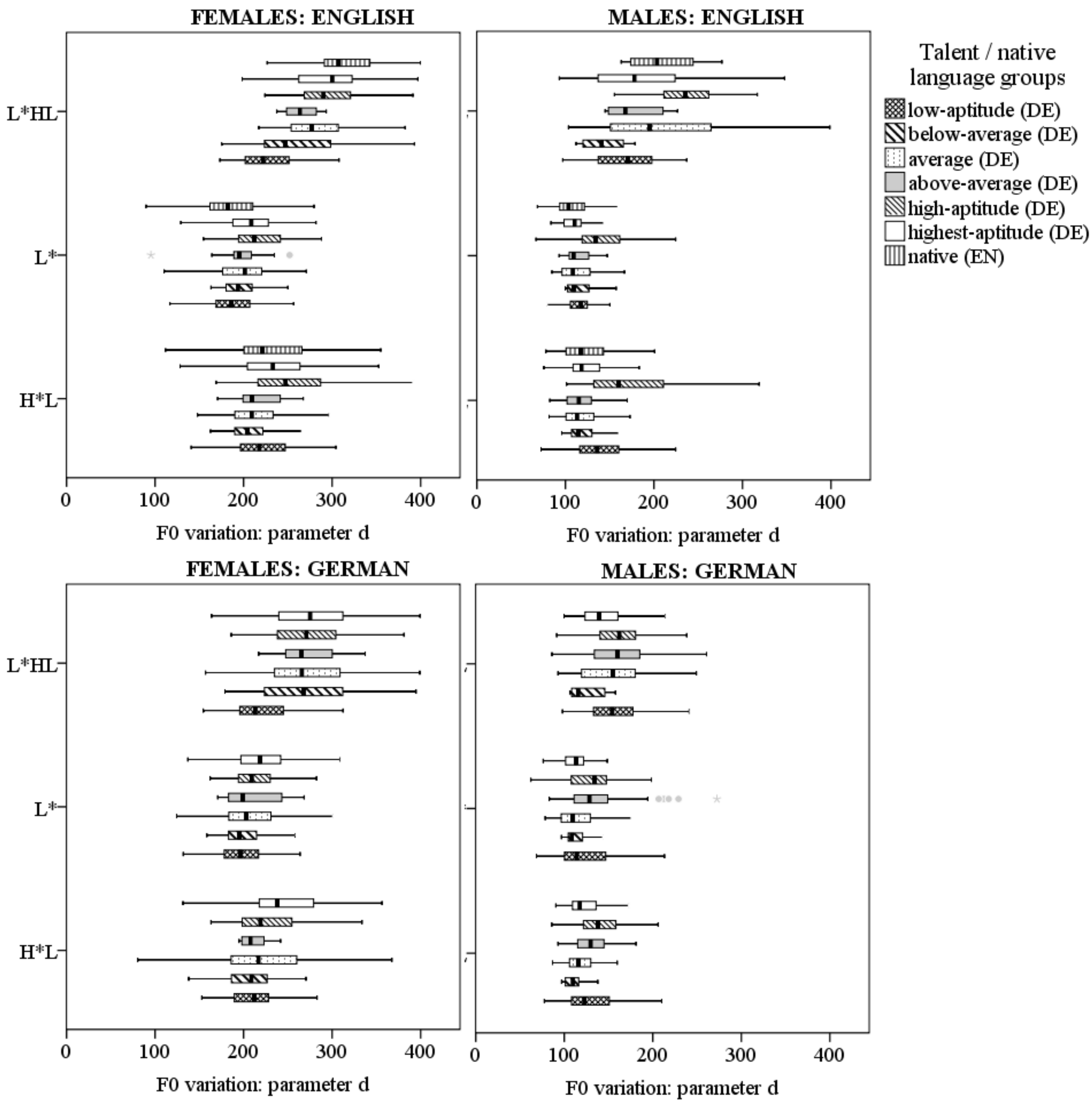


Figure 7.10: d parameter variation in semi-frequent accents (L, H*L, L*HL) with regard to the F0 level in Hz in all the aptitude groups in English and German by gender – male vs. female realizations*

Thus, the most pronounced of such discrepancies involved the above-average group in the two languages under investigation: the values were much more scattered around the mean than the tokens pertaining to the remaining groups.

Moreover, the mean values tended to occupy a lower position than the corresponding targets in the other groups. In German, we could also discover another cluster with a relatively high degree of variability of the *b* parameter – the highest-aptitude, average and below-average groups. The English data would only single out the below-average group in this respect.

The non-uniform distribution most evidently affected the results of the ANOVA tests, which attested to the statistical significance of inter-group differences (p=.000). No Scheffé contrasts were nevertheless significant (p>.6)[39].

After investigating the general tendencies of *b* parameter distribution, we proceeded with the gender-specific investigation.

Gender-specific analysis

Falling-rising tones L*HL

The female samples of the rising-falling L*HL accents thereby were marked by similar degrees of variability across the groups in both languages under investigation. The only exception to this pattern was presented by the aberrant performance of the above-average female speaker with visibly more variable values in English. Also worthy of mentioning was the recurring tendency of L2 speakers to produce their German utterances with more variability in parameter *b* than the corresponding English-language tokens.

This fact was generally supported by the male data as well, with the exception of the low-aptitude group, who, vice versa, produced their targets with less variability in German. The latter subjects also presented the only deviation from the overall uniform variability pattern in German male speech. The same was, but less conspicuously, true of the English-language realizations.

ANOVAs revealed no significant outcome for the male speakers in both languages and for the female subjects in English (p>.1). Cross-group comparisons in German female speech were, on the contrary, found to be significant (p=.019). Nevertheless none of the Scheffé tests pointed out to any significant results (p>.2).

39 Figure 21 of the Appendix, as well as Tables 15-16 provide a detailed insight into the distribution of L*, L*HL and H*L accents in the experimental corpus.

Simple low targets L*

Next we looked at the alignment of *F0* peaks in L* accents. Thus, the female realizations were identified by similar degrees of variability in both English and German. A slight departure from this tendency was found in English in the behavior of the above-average speaker with a less variable sample and the high-aptitude informants with a marginally greater degree of variability than the remaining groups. As for the mean values, they were also in line across the groups and only fluctuated within around 0.2 standard deviations.

Interesting results were obtained for the male data in this experimental condition. Here, the above-average speakers aligned their *F0* peaks in the most variable way in English, considerably differing from the corresponding samples in the other groups. Also visibly more variable than those pertaining to the remaining groups was the sample produced by the high-aptitude subjects. In the German part of the corpus, male speakers exhibited more agreement with respect to variability: only average and below-average subjects had slightly less variation in their realizations. Mean values lay on more or less the same level in all groups in both English and German.

This observation was supported by the ANOVA outcome to a considerable degree: the only significant result was found in the male productions in English (p=.005). Yet again, no Scheffé tests yielded any statistical significance (p>.05).

Falling tones H*L

The falling H*L tones appeared to present a rather heterogeneous picture with respect to the distribution of parameter *b*. In English female speech, for instance, the low-aptitude, below-average and, most of all, above-average speakers could be singled out for their greater degree of variability in parameter *b* than the remaining groups. In German, on the other hand, the above-average female subject preserved the same variability pattern, whereby the other two groups in question did not depart from most other subjects in this sense. These did not, however, include the high-aptitude informants, whose values demonstrated the smallest degree of variability in this condition.

Investigating male speech, we also found quite some variability mismatches between the groups. The English speech of high-aptitude, above-average and average test persons was far more variable with regard to the alignment of the *F0* peaks in H*L accents than the other group samples. The same was, but partially, evident from the German-language realizations. Only in this condition the above-average and average speakers preserved their high variability degree. The low-aptitude group, on the other hand, presented the most mean-concentrated values. Additionally, the tokens produced by the above-average group tended to be produced on a visibly lower mean level than those produced by the rest of the subjects.

This difference might have affected the ANOVA results, which were significant in all of the conditions for H*L accents presented above ($p<.03$). Several Scheffé tests also produced significant results. However, they did appear to suggest any trends in the data[40].

Parameter *a2*

The PaIntE parameter *a2* renders the velocity of the falling sigmoid in a given *F0* curve. It is therefore appears to be an important structural element of the falling H*L and rising-falling L*HL accents, but not the simple low targets L*.

General corpus analysis

The initial observation regarding the distribution of this parameter in both of the speech accents in question allowed us to conclude that there was generally more variability in the English data than in German. Apart from that, there appeared to be more agreement as to the degrees of variability in L*HL events as opposed to the falling tones.

Rising-falling tones L*HL

Thus, the degrees of variability of rise-falls across the groups in each of the languages were almost equal, as were the mean values. Unsurprisingly, this uniformity resulted in the absence of significant contrasts in the ANOVA tests ($p>.5$). Concurring with this finding, Scheffé tests did not yield any significant outcome either ($p>.5$).

Falling tones H*L

The situation with the H*L accents was a little less homogeneous: whereby the English tokens exhibited similar degrees of variation in all groups, the German data presented some cross-group mismatches. More specifically, the low-aptitude subjects demonstrated a clearly more variable sample than the remaining groups; above-average speakers, on the other hand, had the smallest degree of variability in their values, coupled with an abnormally low mean value.

ANOVAs suggested that the inter-group differences were significant ($p<.05$) in both languages, possibly prompted by the outlying performance of the groups mentioned above. Nevertheless no Scheffé contrasts attested to inter-group significance ($p>.2$)[41].

40 Figure 22 and Tables 15-16 of the Appendix are to be referred to for an illustration of cross-gender peculiarities in parameter *b* of L*, L*HL and H*L accents.

Gender-specific analysis

When we analyzed the female and male parts of the corpus separately, we discovered some additional discrepancies as compared to the investigation of the entire corpus.

Rising-falling tones L*HL

In L*HL accents, this was more characteristic of the English-language samples, possibly due to the absence of the rising-falling accent category in the English ToBI per se. More conspicuous in this respect were the realizations of male speakers. The highest-aptitude and average groups were singled out for their greater degree of variability, and the below-average subject as well as, especially, the native informants produced the values to a much greater extent concentrated around the mean. The remaining speakers realized their *a2* tokens in a much similar way with regard to variability. Unlike their male counterpart group, the female native speakers produced the most variable sample of all female informants in English.

The German-language realizations of both the female and male subjects were marked by comparable variability degrees.

Most ANOVA tests brought non-significant outcomes (p>.4), except for the English male-speech condition (p=.006). None of the post-hoc tests attested to statistically non-significant differences between the groups (p>.6)[42].

Falling tones H*L

The falling tones were rather evenly distributed in English female speech, with both the degrees of variability and mean levels being comparable in most groups. With regard to the former parameter, the above-average speakers departed from the pattern, producing a less variable sample and lower mean values (~9 vs. ~13 on average). Notably, the below-average female speakers had similarly low mean values. The native speakers, on the contrary, presented a sample with the highest mean values (~16).

In their L1 speech, the German groups also displayed a few cross-group differences. In particular, the above-average and low-aptitude groups rendered the most variable samples, which in the case with the latter group was also coupled with a higher mean level (~14 vs. ~10-12). The below-average female subjects, on the other hand, had the lowest mean values (~9).

41 Figure 23 and Tables 17-18 of the Appendix refer to the findings in the *a2* parameter values in L*HL and H*L accents.
42 Figure 24 and Tables 17-18 of the Appendix illustrate the results of the gender-specific analysis of *a2* tokens in L*HL and H*L accents.

Most mean values in English male speech fluctuated between ~12 and ~16, with the exception of the above-average group with a lower mean (~9). The degrees of variability were, for the most part, comparable in this condition. Only the low-aptitude group exhibited a greater spreading of values around the mean.

The latter group was marked by similar performance in German. However, in this condition the highest-aptitude group presented a similarly variable sample. In the remaining groups, the degrees of variability were lower to an almost equal extent, and so were most mean values, except for the above-average subject realizing the *a2* parameter targets on a lower mean level (~5 vs. ~10-11 in the remaining groups).

None of the above distinctions were statistically significant in the course of the ANOVA tests and the Scheffé corrections (p>.05).

Parameter *c2*

The final parameter under investigation was the amplitude of the falling sigmoid – *c2*. As with *a2*, we restricted ourselves to the description of the L*HL and H*L accents in this respect, since the production of simple low targets L* does not presuppose any significant downward pitch movement.

General corpus analysis

Rising-falling tones L*HL

On the whole, the distribution of parameter *c1* in the rising-falling tones presented a rather heterogeneous variability picture, with the only exceptional condition being L*HL accents in German. Here, the degrees of variability were almost equal in all groups, but for the below-average subjects with slightly more variable sample.

L*HL accents in English singled out the cluster of the native, highest-aptitude and high-aptitude groups as the speakers having realized their falling sigmoids in the most variable way, which was followed by the performance of average and below-average informants. Finally, the smallest degree of variability was discovered in the low-aptitude and above-average group realizations.

In English, the differences between the groups were found to be significant due to the fluctuations in mean values (p=.001), whereby the high-aptitude subjects demonstrated the highest mean level of around 70 Hz. The values in the other groups were lower (~50-60 Hz). However, only the contrast between the high-aptitude and low-aptitude speakers was confirmed to be significant in the course of the Scheffé procedures (p=.010).

For the German part of the data, ANOVA did not yield any significant results (p=.272), nor did the post-hoc Scheffé tests (p>.05).

Falling tones H*L

The first observation we could make about the falling tones regarded the increased degrees of variability in the average, the high-aptitude and, to the greatest extent, highest-aptitude speakers' samples in German, which departed from the smaller degrees of variability in the remaining groups, of which the above-average subjects were identified by the smallest degree of variability. In their English productions, the German informants exhibited a somewhat similar variation behavior as in their L1. Thus, the highest-aptitude and high-aptitude subjects were singled out due to their high degree of variability. The average group, on the other hand, clustered with the low-aptitude group and the native test persons. The above-average and below-average speakers found themselves at the lower end of the variability scale in this condition.

ANOVAs stated the significance of inter-group differences in the falling tones in parameter *c2* (p=.000). A more detailed cross-group comparison rendered by the Scheffé tests attested to one consistent pattern: in English the above-average group appeared to present a deviant sample differing from all the rest of the group realizations (p<0.01)[43].

At the final stage of pitch accent analysis, the *c2* parameter tokens were analyzed with regard to the corresponding gender samples for further analysis.

Gender-specific analysis

Rising-falling tones L*HL

As in several previous conditions, it was possible to position the native, highest-aptitude and high-aptitude female groups in English into one cluster of speakers with the greatest degrees of variability, whereby the latter subjects demonstrated the most variable sample of all. These three groups were, in turn, followed by the average and below-average subjects in performance. The remaining two groups had least variation in their values. The German data, on the contrary, appeared to follow a rather uniform variation pattern in the female productions. Here, the values of the high-aptitude informants only had slightly more and the below-average test persons – slightly less variability than in the other groups.

The mean *F0* levels in English differed to a substantial extent across the female groups, whereby the native and highest-aptitude speakers demonstrated the highest values (~100 Hz and ~80 Hz,

43 Figure 25 and Tables 19-20 of the Appendix render the general corpus information of parameter *c1* distribution in L*HL and H*L accents.

respectively). The mean values in the remaining female groups fluctuated between ~40 and ~60 Hz. In the German part of the corpus, most aptitude groups realized their *c2* parameter tokens at ~50-60 Hz. The low-aptitude speakers, however, rendered an aberrantly lower (~25 Hz) and the below-average group – a higher mean *F0* level (~70 Hz).

The realization of L*HL accents by L1 German male subjects in English did not contain any substantial mismatches as to the degrees of variability. As far as the mean levels are concerned, the highest-aptitude and the below-average test person produced lower mean levels than the remaining subject pool (~20-40 Hz vs. ~50 Hz); high-aptitude male speakers, vice versa, demonstrated abnormally high average values (~70 Hz).

German male speech exhibited decreasing degrees of variation from the above-average, average and low-aptitude; to the below-average speaker with the greatest dispersion of values. The only observation concerning the mean *F0* levels was the fact that the above-average subjects were marked by the highest mean level (~50 Hz).

All ANOVA tests produced significant results ($p<.01$). Significant statistical comparisons in Scheffé tests were not systematic.

Falling tones H*L

Turning the focus of attention to the falling H*L accents, we were able to associate the highest-aptitude, high-aptitude and native female informants in English, both due to their similarly greater degrees of variability and higher mean pitch than in the other groups (~60-70 Hz vs. ~30-40 Hz), with the exception of the low-aptitude speakers with similarly high mean values. Contrary to that, below-average female subjects produced the least variable sample in this condition. The remaining female speakers took an intermediate position in this respect.

The distribution of *c2* parameter values in H*L accents in German female speech differed from the corresponding pattern in male samples. Whereas the highest-aptitude and high-aptitude subjects preserved their high degrees of variation and low-aptitude and average speakers retained the mid variability level, the above-average group displayed the smallest degree of variability in this condition. The below-average female informants also performed differently by increasing their degree of variability as opposed to English.

The main discrepancies between the groups in English male speech involved the high-aptitude and low-aptitude groups with higher degrees of variation and higher mean levels (~40 vs. ~30 Hz). The below-average subject once again performed aberrantly with the lowest degree of variation. In German only two groups could be singled out as falling out of the otherwise uniform pattern: the below-average and average subjects had visibly less variability in their values than the other groups in this condition.

All ANOVA tests were significant ($p<.01$). Nevertheless no cross-group differences in German were found to be significant, according to the Scheffé routine. In English, the high-aptitude male group with their high mean values was statistically opposed to almost all groups ($p=.000$), except for the low-aptitude test persons ($p=.505$)[44].

7.1.2 Boundary tones

The investigation of boundary tones was limited in the current study to the exploration of parameter *d* due to the fact that pitch peak level appeared to be the most characteristic for phrase edge events, and the definition of the remaining PaIntE parameters on boundary tone level did not appear to be possible.

In general, we described the peculiarities of distribution of the following boundary tone categories: L-L% (full low), H-H% (full high), % (default), L- (intermediate low). These units underwent a detailed investigation for they were detected as being the most frequent categories at the stage of phonological analysis.

General corpus analysis

Consequently, a standard procedure was applied to their investigation and description: first the entire experimental corpus was taken as a whole; next the factors of gender and speech type were analyzed separately; finally, the combined analysis of speech type / gender samples concluded the procedure.

Full low boundary tones L-L%

We first looked into the most frequent category in the corpus – the low intonation phrase boundary tones – L-L%. In both languages under investigation, the above-average group demonstrated conspicuous performance by producing visibly lower mean pitch levels than the remaining subjects (~110 Hz vs. ~180-190 Hz). Notably, in English the native subjects approached this outlying mean level: they pronounced their prominent syllables only ~10 Hz higher in pitch on average than the above-average group.

These idiosyncrasies in pitch levels were supported by statistical analysis. Thus, Scheffé tests stated the difference of the above-average group sample from all the other ones in German ($p<.01$). A similar situation was observed in English with these speakers, where they were only

---

44 Figure 26 and Tables 19-20 of the Appendix render the gender-specific results of parameter *c1* distribution in L*HL and H*L accents.

found to be statistically similar to the native subjects (p=.337). Post hoc tests also attested to the significant differences between the latter group and high-aptitude, average and below-average test persons (p<.01). Interestingly, the low-aptitude speakers also deviated from the majority of samples in parameter *d* of L-L% events, despite the fact that their mean values were only marginally lower than in most groups, without account of the outlying above-average and native speakers. All these comparisons yielded statistically significant p-values with the exception of the contrasts against below-average informants in German (p=.146) and average test persons in English (p=.115).On the whole, the inter-group differences were found to be significant in the ANOVA tests (p=.000) for both the English and the German segments of the corpus.

With respect to the degrees of variability in L-L% events, one ought to point out that all samples were comparable in both language conditions. Only the highest-aptitude informants had slightly more variation in values than the rest of the subject pool. Below-average test persons, on the other hand, were singled out for their lowest degree of variability, which was especially evident in English.

High full boundary tones H-H%

Unlike the low full boundaries, the H-H% events did appear to suggest certain cross-group peculiarities in degrees of variability. First of all, the highest-aptitude informants realized their full high boundaries in the most variable way in both languages under investigation. In German, such performance was replicated by the high-aptitude subjects, who then, however, dropped their variation level in the English-language realizations. Native speakers, on the other hand, had a similar degree of variability of H-H% tokens in English as the highest-aptitude group. A considerably high degree of variability in parameter *d* of the full high boundary targets was typical of the average group, which almost equalled the corresponding values observed for the high-aptitude test persons. Finally, the above-average, below-average and low-aptitude speakers had the most mean-concentrated values.

With respect to the mean pitch, we could draw a similar observation about the above-average subjects in H-H% tones as the one noticed for the full low boundaries: this group was characterized by the lowest mean *F0* level of all groups (~150 Hz vs. ~190-210 in English; ~180 Hz vs. 200-230 in German). Other than this peculiarity, most group samples pertained to comparable *F0* levels. The only two exceptions in that sense were presented by the below-average group in German with a slightly higher than average pitch (~230 Hz vs. ~210-220 Hz) and the native speakers in English, whose mean *F0* values were, vice versa, slightly lower than those produced by the other informants (~190 Hz vs. ~200-210 Hz).

These differences in pitch levels were found to be statistically significant as a result of the ANOVA tests (p=.000). More exact cross-group details were rendered by the Scheffé post hoc comparisons. Thus, a separation of the above-average group in statistical terms was established as opposed to all groups in English (p<.05) and most other speakers in German, except for the low-aptitude subjects (p=.119). The latter also appeared to demonstrate statistically conspicuous behaviour in German: the contrasts with the high-aptitude, average and below-average group samples were significant (p<.03). The English data segment singled out native-group productions as being statistically deviant from most samples, but for the highest-aptitude and low-aptitude groups (p>.21), exhibiting closest mean *F0* values.

Default boundary tones %

For the default full boundaries %, we discovered generally similar, but less pronounced, tendencies in degrees of variability as the ones registered for the H-H% tokens. Here, the highest-aptitude group once again produced the most variable values. However, in German, almost identical performance was demonstrated by the average group; and in English – by the native speakers, whereby the above-average, average and low-aptitude test persons closely followed. The rest of the samples had less variation in PaIntE parameter *d*.

The mean pitch level distribution and degrees of variability in default boundaries also reminded of that in H-H% events. More precisely, the above-average subjects were marked by the lowest *F0* levels in this condition as well (~120 Hz vs. ~180-200 Hz in German; ~120 Hz vs. ~150-200 Hz in English). In the English part of the corpus the native speakers also followed the tendency for lower mean pitch (~150 Hz), which was most closely followed by the high-aptitude and low-aptitude groups (~170 Hz).

The highest-aptitude informants continued to demonstrate the highest degree of variability in boundary tone tokens, namely, in both languages at issue. In German, average speakers exhibited almost equal variability of values, and in the English part of the corpus they approached it with a slightly less variable sample. The remaining groups had less variability in % boundary tone values in German, whereby the samples rendered by the above-average and below-average groups were more centred around the mean. In English, all L1 German groups were characterized by similar degrees of variability, except for the above-mentioned below-average informants.

Pitch level mismatches and the general samples differences were significant (ANOVA, p=.000). A distance of the above-average group from the other subjects due to their low mean *F0* level was established statistically in the Scheffé tests for both English and German: in the latter language all comparisons were significant (p<.01); in English – with the exception of the juxtaposition with the native group (p=.704). The relatively low mean pitch values of the latter speakers contrasted them statistically with the high-aptitude, average and below-average subjects (p<.01).

Intermediate low boundaries L-

The final type of boundary tones, i.e. L-, yielded a few cross-group discrepancies, but only one major observation in degrees of variability. Almost identical behaviour, or highest variability, was discovered in parameter *d* for the below-average and highest-aptitude groups in German and English speech. All the other speakers had lower similar degrees of variation in this condition.

With regard to the mean pitch levels, the above-average subjects demonstrated the lowest values (~140 Hz vs. ~200 Hz in German; ~140 Hz vs. ~170-200 in English), but in German the same level was detected for the low-aptitude informants; and in English – for the native subjects, followed by the low-aptitude group.

ANOVA results lay below the significance level in the general comparison ($p<.01$). However, the only significant contrast in the Scheffé post hoc tests constituted the comparison of the native and the average groups ($p=.001$).

Figure 7.11 of the current Section illustrates the distribution of *d* parameter values in all of the investigated boundary tones. Statistical information of the post hoc Scheffé tests is to be found in the Appendix: Tables 21-28.

After completing the analysis of the entire corpus, we undertook the gender-specific comparison.

Gender-specific analysis

At the first glance, the female data appeared to present a more homogeneous picture both with respect to the degrees of variability and the mean *F0* levels than male realizations.

Full low boundary tones L-L%

In the low full boundaries L-L% of the German-language part of the corpus the differences between the samples as to the variability degrees were minor; so were the mean pitch values.

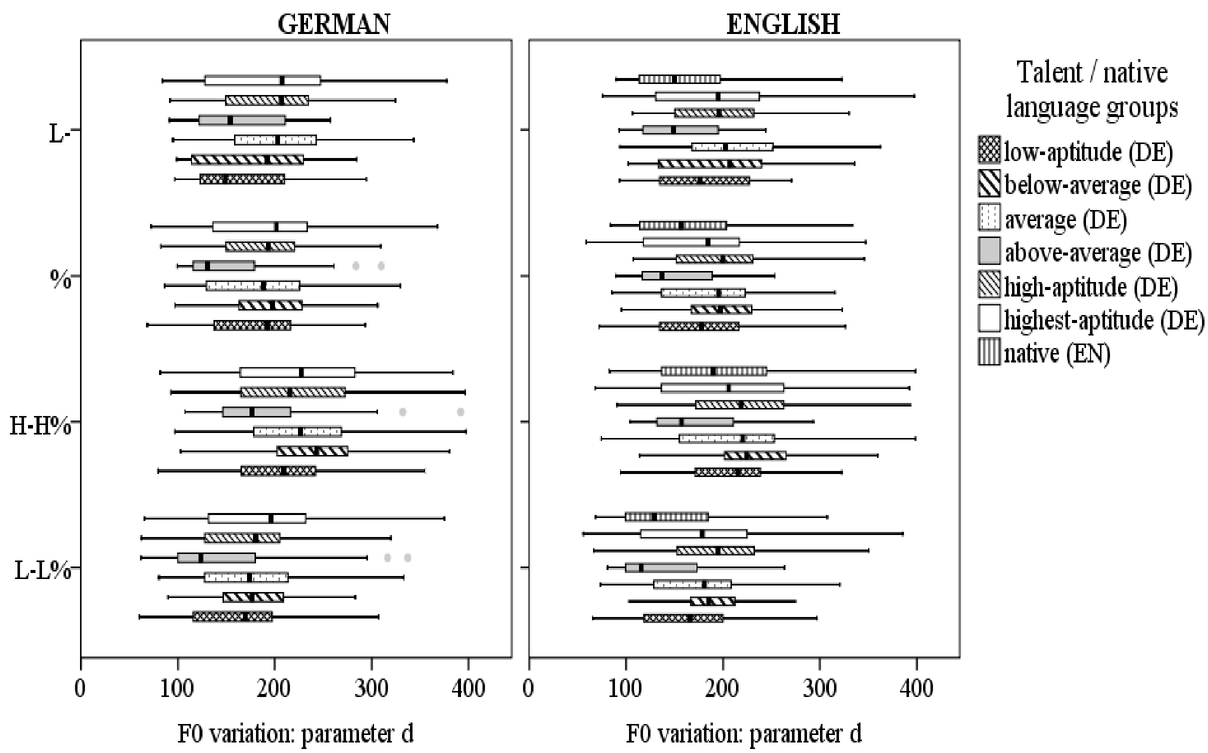


Figure 7.11: d parameter variation in boundary tones (L-L%, H-H%, %, L-) with regard to the F0 level in Hz in all the aptitude groups in English and German

The only group departing from the given pattern were the highest-aptitude informants with a somewhat higher than average pitch (~210 Hz vs. ~190 Hz).

The English-language utterances, in their turn, were an example of aptitude / variability clustering, observed for a number of previous experimental conditions: native, highest-aptitude and high-aptitude female speakers had samples of almost equal variability. The remaining groups had visibly less variable productions of L-L% boundary tones. Still another recurring peculiarity was the mean *F0* level in the highest-aptitude and high-aptitude groups, overhitting the average targets in the other groups by about 20 Hz (with mean *F0* values of ~250 Hz), including the native speakers (~220-230 Hz).

The distribution of values by the male groups in German was relatively regular. The only outlier in relation to variability was, as in several earlier cases, the below-average subject, who produced a sample rather tightly concentrated around the mean. The only other deviation from the general trend was presented by the high-aptitude informants. In particular, these speakers had slightly more variability in their L-L% tokens, as well as realized them on a considerably higher mean pitch than the remaining groups (~135 Hz vs. ~100-110 Hz).

High-aptitude male test persons were also singled out for their conspicuous performance in English: not only was the sample to a considerable extent more variable than in the other male groups, the mean *F0* values were also located much higher on the scale (~150 Hz vs. ~100-120 Hz). The remaining groups realized full low boundaries in a less variable way.

All of these differences taken as a whole were treated as significant in the ANOVA procedure – in both languages and for both female and male informants (p=.000). The inter-group differences described above also found reflection in the Scheffé tests. A contrast was therefore confirmed in German between the highest-aptitude female group, on the one hand, and average, below-average and low-aptitude subjects, on the other (p<.02), apparently, due to the higher mean *F0* of the former. The deviation of the high-aptitude male test persons in German was also proven by the Scheffé tests (p<.05).

The English samples also provided evidence for the empirical observations regarding the distribution of parameter d in L-L% boundary tones. In the female productions, the highest-aptitude and high-aptitude speakers were separated statistically from the rest of the subject pool (p<.01), having exhibited higher than average pitch in this condition. Only the comparison of these two groups against one another was not significant (p>.05). Predictably, the high-aptitude male group, with their high *F0* values, presented an example of consistent statistical deviation from the remaining male groups (p=.000).

High full boundary tones H-H%

The analysis of the full high boundary tones allowed us to conclude that both the female and male productions in German were relatively uniform in degrees of variability. Only two groups departed from the pattern in each case: of the female subjects, highest-aptitude test persons had slightly more variation in values, whereas the above-average informant – marginally less; low-aptitude and highest-aptitude male test persons exhibited a lower degree of variability than the remaining speakers in male speech.

There were quite some mismatches in the mean *F0* levels on the cross-group level genderwise.

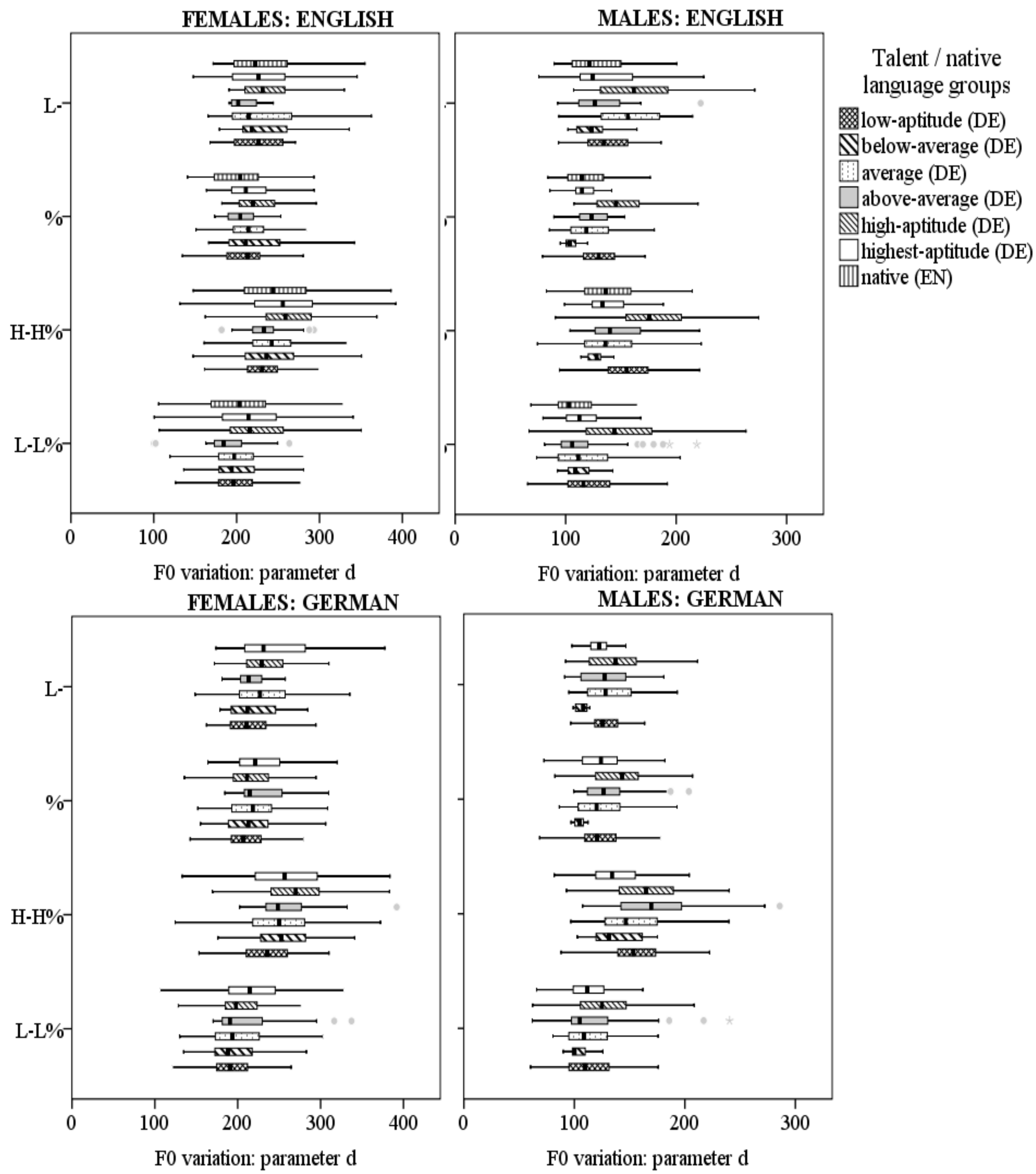


Figure 7.12: d parameter variation in boundary tones (L-L%, H-H%, %, L-) with regard to the F0 level in Hz in all the aptitude groups in English and German by gender – male versus female realizations

Looking at the German-language data, we could see that in the female utterances, the highest-aptitude, above-average, average and below-average subjects all realized their H-H% tokens on an almost equal mean pitch (~250 Hz). The high-aptitude group deviated hereby by exhibiting higher (~270 Hz) and the low-aptitude informants – lower (~230 Hz) mean pitch.

The male part of the corpus could be split into three clusters based on the mean *F0*: highest-aptitude and below-average speakers with the lowest values (~120 Hz); the average and low-aptitude subjects on an intermediate-level position (~150 Hz); highest-level values demonstrated by the high-aptitude and above-average informants (~170 Hz).

The English part of the corpus presented more cross-group discrepancies than the German speech. This was most evident from the utterances pertaining to the male subjects. The high-aptitude informants, in particular, appeared to be the most remarkable outlier with regard to the mean pitch (~170 Hz vs. 110-120 Hz in most other male groups), as well as their increased degree of variability. Still another outlying occurrence was the low mean *F0* level reached by the low-aptitude speakers (~150 Hz), who differed both from the above-mentioned high-aptitude group and the other male subjects. The degrees of variability were comparable in most groups, excluding the highest-aptitude and, especially, the below-average subject, who realized the H-H% tokens less variably.

Next we examined the female realizations in English, where similar clustering as in several previous conditions was discovered, e.g. in the whole corpus of low full boundaries: the native and the highest-aptitude groups shared equal degrees of variability. The high-aptitude subjects, unlike within the L-L% category, did not produce a comparably variable sample for the L-L% targets. Instead, these speakers could be positioned with the below-average and the average informants on this account – by demonstrating less variation in H-H% target values. The above-average female speaker performed aberrantly, as the values produced by this informant contained the smallest degree of variability.

With regard to the mean *F0* parameter, most groups delivered similarly pitched syllables (~220 Hz). However, in the highest-aptitude and high-aptitude test persons' realizations the mean pitch level was elevated (~240 Hz) – an observation drawn for several previous conditions.

The above-mentioned pitch differences in *F0* levels were significant (ANOVA, p=.000). The low mean *F0* in the low-aptitude female group in German separated these speakers from most other groups (p<.01), except for the above-average test persons (p=.236).

Male speakers of highest pronunciation aptitude, in their L1 utterances, appeared to deviate quite substantially from the other speakers as well (p<.03), whereby they would only agree with the below-average group in this sense (p=.020). In English, the high-aptitude group of both genders demonstrated conspicuous behaviour: the male speakers with their abnormally high mean *F0* level also caused their sample to be highly statistically different from all the investigated groups without exception (*p*=.000); most contrasts for the female informants were also significant, not taking into account the highest-aptitude group (*p*=.656). The remaining significant tests were not systematic.

Default boundary tones %

The default boundaries %, a common German boundary category, not typically characteristic of English, revealed uniform degrees of variability in most German groups of both genders in their L1 utterances, and in the case with the female subjects also an agreement with respect to the mean pitch levels (~200 Hz). Male informants deviated from this general homogeneity in a number of aspects. First of all, the below-average subject demonstrated the lowest degree of variation coupled with the lowest mean *F0* values (~90 Hz vs. ~110-120 Hz). The other group with outstanding performance, the high-aptitude subjects, were only singled out due to their idiosyncratically high mean pitch (~140 Hz vs. ~110-120).

The English part of the data also unveiled quite a number of cross-group mismatches. As for the degrees of variability, in most groups the default boundary tokens were less variable than in German. However, the spreading of values in the native and the below-average female groups was considerable. Apart from this deviation, the female samples were quite uniform, which also applied to the mean *F0* levels (~210 Hz in all groups).

We further analyzed male speech in the English-language condition. A greater degree of variability and a visibly higher pitch appeared to be characteristic of the male high-aptitude group, as similar performance had been discovered in several previous experimental conditions and repeated for the distribution of the default boundary values. Indeed, the distance between the mean *F0* level produced by these subjects as opposed to the remaining groups was substantial (~150 Hz vs. ~120-130 Hz). The other male groups rendered samples of lower variability in this condition, which was especially evident from the below-average subject's performance: the sample was characterized by a very low degree of variability, as well as the lowest mean pitch in this condition (~90 Hz).

Logically following the empirical observation on the general differences in pitch levels, the ANOVA results were significant for both males and females in the two languages at issue (*p*=.000). The specific cross-group comparisons of the Scheffé procedure also reflected the remarks expressed above. The greater consistency and homogeneity in female values only triggered several statistically significant contrasts (*p*<.05), which did not, however, suggest any systematic cross-group mismatches.

As expected for the male speakers in German, two groups could be singled out due to their idiosyncracy in mean *F0* values of parameter *d* of the default boundary tones – the high-aptitude and the below-average speakers. Of all the comparisons, only the one involving the high-aptitude and above-average groups lay above the significance level (p=.350). In the English speech, the high-aptitude informants were the sole group to differ statistically from all the other subjects (p=.000).

Intermediate low boundary tones L-

The final edge tone category, the low intermediate boundaries L-, revealed some further cross-gender and cross-group differences. Thus, the female subjects were generally found to be more consistent when producing their L- tokens. The greatest departure from this consistency was in this condition the performance of the above-average female speaker with a visibly lower degree of variability than in the other groups in both languages. And in English, this subject also had the lowest mean *F0* level (~190 Hz vs. ~200-210 Hz). In the L1 productions the mean values were slightly higher, and in line with those of the below-average and low-aptitude speakers (~200 Hz). The average, high-aptitude and highest-aptitude test persons produced their L- targets on a slightly higher mean pitch (~210 Hz). Other than the aberrant performance of the above-average female speaker, the other female groups performed in a generally comparable way with regard to variability.

When we looked at the male data in the L- category, we could once again pinpoint the below-average subject for the lowest degree of variability and the lowest mean *F0* in German (~100 Hz vs. ~130 Hz). In the English utterances the below-average subject could be grouped with the native, highest-aptitude and above-average informants for exhibiting similar mean pitch values (~110 Hz).

Further on, we could draw some observations from the other male groups' behaviour. In German, for instance, the degree of variability in the highest-aptitude and low-aptitude groups was almost equally low as in the below-average speaker's productions. The other test persons performed in a similar way in that respect. The high-aptitude speakers, however, demonstrated slightly higher than average mean *F0* values (~140 Hz vs. ~130 Hz).

In English, this contrast was visibly greater, but here the average group produced similarly high mean values (~150 Hz vs. ~120-130 Hz). The degree of variability was highest in the high-aptitude group.

Most of the ANOVA tests stated the insignificance of cross-group differences, with the exception of the male speakers in English (p=.000). Further on, only two post-hoc contrasts were significant: the native male subjects vs. the high-aptitude and below-average male groups (p<.03).

Figure 7.12 reports on gender-specific findings in parameter *d* of boundary tones.

Corpus analysis by speech type

The initial examination of the parameter *d* distribution by speech type revealed some clear differences between the groups both with regard to the degrees of variability and the mean *F0* levels.

Full low boundary tones L-L%

The German-language productions of the L-L% boundary tones did not contain many cross-group mismatches. Thus, in read speech, only the above-average sample had conspicuously low mean *F0* values (~140 Hz) against similarly higher levels in the remaining groups (~180-190 Hz). The degrees of variability were almost equal in all groups in German read utterances. German spontaneous speech yielded more differences between the aptitude groups. Firstly, high-aptitude informants demonstrated here not only a very low mean pitch (~100 Hz vs. ~180-210 Hz in the other groups, maximal values being found within the highest-aptitude informants) – they also had the least variable sample. Almost an equally low degree of variability was exhibited by below-average group, similarly to their read-speech realizations.

In English read utterances, the aptitude groups behaved to a large extent similarly to their L1 performance, yet with some deviations. More specifically, the mean *F0* levels in all German groups for the most part replicated the L1 pattern – with the above-average group exhibiting the lowest values (~120 Hz) and the remaining groups generally keeping to similar levels (~180-190 Hz). However, a slightly higher mean pitch was typical of the high-aptitude test persons in this condition (~200 Hz). The mean values produced by the native speakers were adjacent to those of the above-average group, i.e. relatively low.

The degrees of variability in English read speech were comparable across the samples. However, the highest-aptitude subjects had quite visibly more and the below-average speakers – marginally less – spreading of values around the mean.

Spontaneously produced utterances carrying L-L% edge tones in English departed from their German-language counterparts. It was, first of all, evident from the mean *F0* levels. The above-average subjects thereby established the pitch minimum in this condition (~100 Hz) and below-average speakers were characterized by the highest mean *F0* values (~210 Hz). The pitch levels in other groups were located between these so-called boundaries, falling into three relative clusters: the native and highest-aptitude groups with the lowest levels (~130 Hz); the low-aptitude informants (~160 Hz) on mid pitch; as well as the average and high-aptitude speakers with the highest values (~ 190 Hz).

As far as the degree of variability in English spontaneous speech were concerned, the above-average and below-average subjects demonstrated the lowest degrees of variation, as opposed to the similarly more variable samples rendered by the remaining groups.

The ANOVA test attested to the significance of inter-group differences in L-L% boundary tone tokens in both read and spontaneous speech in the two investigated languages (p=.000). Further on, we looked at the cross-group mismatches more specifically with the application of the Scheffé post-hoc tests. Thus, we attempted to assess the consistently very low *F0* values produced by the above-average speakers in all of the experimental conditions. The results were predominantly significant in both languages in the two speech types (p<.04). The exceptions to the tendency included the contrasts against the low-aptitude group in German read speech (p=.432), native speakers in both speech types in English (p>.23) and the above-average subjects in English spontaneously produced utterances (p=.150).

Three other groups also produced consistently different samples as opposed to the other speakers: the highest-aptitude informants in German spontaneous speech with their high mean *F0* level (p<.02), in all contrasts, except for the below-average group; high-aptitude subjects in English read speech (p<.03); and native speakers in English read speech with their generally lower than average mean pitch (p<.03) not including the comparison with the above-average test persons. The remaining significant contrasts did not appear to present any systematicity.

High full boundary tones H-H%

Analyzing the high full boundary tones H-H%, we first looked at the read-speech productions, which presented a generally uniform pattern. With regard to variability, in both languages under investigation the highest-aptitude group had samples marked by the greatest degree of variability. The high-aptitude and average groups followed closely with slightly smaller degrees of variation in both their L1 and L2 speech; the native speakers approached the two groups in question in English. Lower degrees of variability were proper to the remaining groups in the two language conditions. Whereas in German the above-average, below-average and low-aptitude speakers were almost indistinguishable as regards variability, the English samples presented a visible mismatch – from a higher degree of variation demonstrated by the above-average informants and the most mean-concentrated values pertaining to the below-average group.

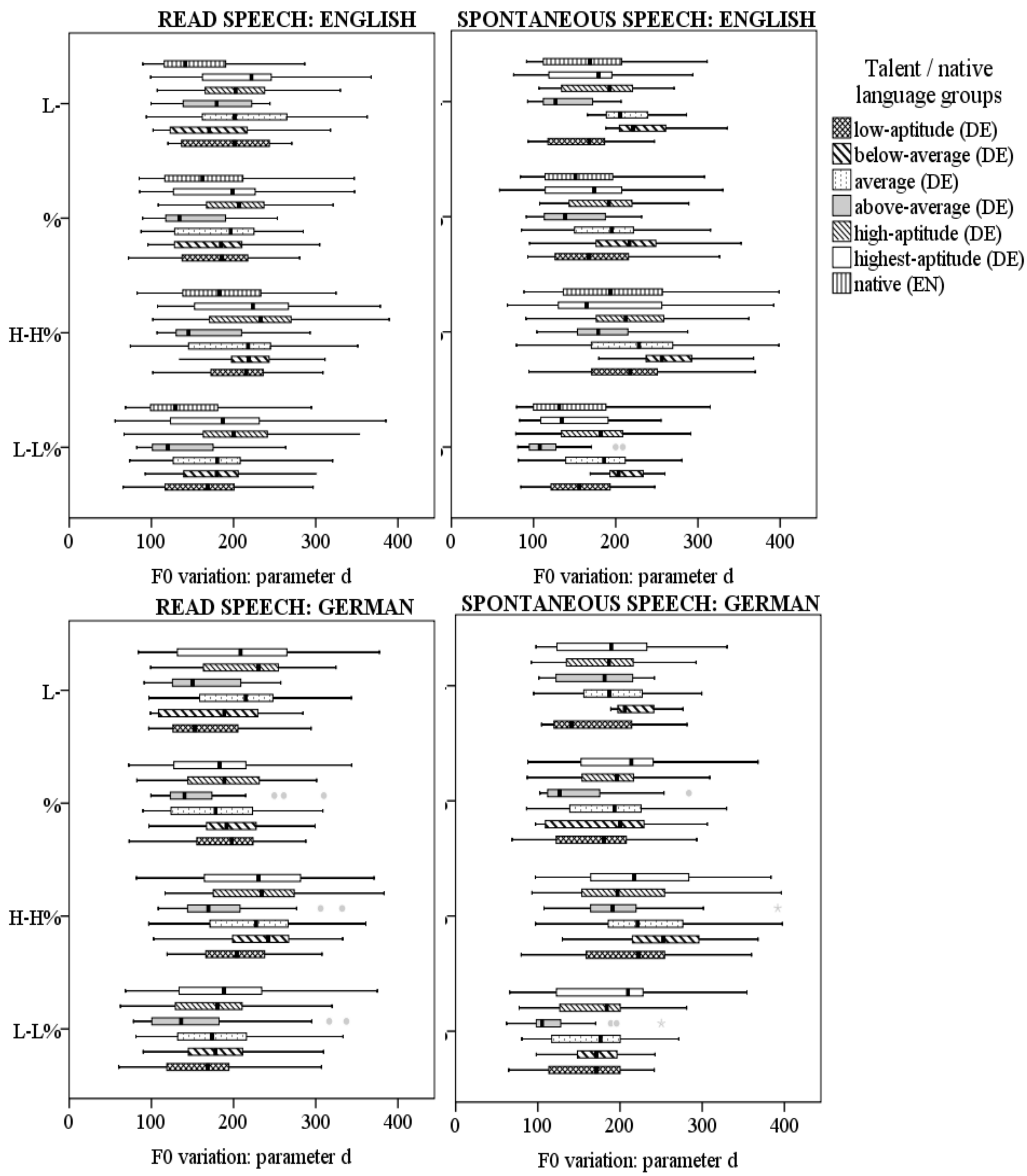


Figure 7.13: d parameter variation in boundary tones (L-L%, H-H%, %, L-) with regard to the F0 level in Hz in all the aptitude groups in English and German by speech type – read versus spontaneous speech realizations

The mean *F0* levels within the H-H% events in read speech presented some stable tendencies, as well as inter-group deviations. As to the consistencies, most groups shared a similar mean pitch of around 220-230 Hz in both English and German.

We could, however, observe some groups departing from this trend, by exhibiting lower mean *F0* values. These groups included the native (~170 Hz) and the above-average informants (~130 Hz) in English; German speech highlighted the low-aptitude (~200 Hz) and the above-average (~170 Hz) informants in this respect.

Spontaneous speech productions of parameter *d* in H-H% boundary tones, in their turn, were reminiscent of the distributions in read speech, especially in German-language productions. Thus, we could observe almost the same degrees of variability in the groups as in the German read speech: highest degree of variation yielded by the highest-aptitude informants; the least variable values were found within the below-average as well as above-average informants.

The pattern in mean *F0* values in German spontaneous speech did not replicate the corresponding read-speech trends. Here, the values presented different clustering: high-aptitude and above-average subjects with the lowest mean pitch (~180-190 Hz); highest-aptitude, average and low-aptitude groups occupying an intermediate position (~210 Hz); finally, the highest mean level rendered by the below-average group (~230 Hz).

During the analysis of English spontaneous speech, we also noticed some similarities in read-speech realizations. We could therefore single out the H-H% targets of the highest-aptitude and native groups with the highest degree of variability. The most mean-concentrated samples, on the other hand, were produced by the above-average and below-average speakers.

The mean *F0* values in English spontaneous speech displayed different performance by the aptitude group, as compared with read speech. The highest level in the below-average group produced the highest *F0* levels (~260 Hz), which was followed by the visibly lower pitch of the high-aptitude, average and low-aptitude informants (~210 Hz). Within the native, above-average and the highest-aptitude subjects the values lay on the lowest level in this condition (~170-190 Hz).

ANOVAs produced significant results for all language and speech type conditions (p=.000). According to the results of the Scheffé procedure, the only group singled out for their consistently different performance in statistical terms were the above-average speakers in German as well English read speech (p=.000). Only the contrasts of these speakers against the low-aptitude group in German and native subjects in English were not significant (p>.15).

Default boundary tones %

The next category under analysis was the default boundary. In both language conditions, we could observe similar mean *F0* level tendencies across the two speech types: the values in most groups fluctuated between ~180 and ~210 Hz. Visibly lower pitch was, contrary to that, discovered in the above-average group realizations in L1 and L2 (~130 Hz). Closest to such low level, native speakers demonstrated almost equally low mean *F0* in English (~140 Hz).

With respect to the degrees of variability, we registered some language-specific trends. Thus, in German, the degree of variation within the highest-aptitude subjects was highest for both speech types, which was closely approached by the high-aptitude and average groups. The remaining German-language read-speech samples demonstrated similarly less variable *d* parameter values in the default boundary tones. In German spontaneous speech, on the other hand, the low-aptitude test persons could be grouped with the highest-aptitude, high-aptitude and average groups, due to their high level of variability in this condition. The above-average speakers' productions had the lowest degree of variability.

The English-language utterances revealed similar clustering in spontaneous speech, as the one described for German above. More specifically, in read speech, the higher than average degree of variability in the highest-aptitude, high-aptitude and average groups was replicated by the native speakers. Spontaneous speech presented a slightly different picture. Here, only the highest-aptitude and native groups demonstrated the greatest scattering of values around the mean *F0* level. The samples rendered by the high-aptitude and average groups were visibly less variable in spontaneously produced utterances. The remaining speakers presented even less variability in the *d* parameter of default boundary tones.

As part of the standard routine, ANOVA tests were carried out to compare the language / aptitude group data subsets, which resulted in *p*-values below the significance level in all of the contrasts ($p=.000$). According to the Scheffé tests, only the above-average group in German spontaneous speech deviated consistently from the other samples due to their low mean *F0* ($p<.03$), except for the contrast with the low-aptitude group ($p=.202$).

Intermediate low boundary tones L-

The low intermediate boundary tokens appeared to exhibit more variability in read speech in all groups, as contrasted with the spontaneous speech productions. Nevertheless the cross-group differences in degrees of variability did not appear to present a clear-cut pattern. In their L1 speech, the German groups had almost equal degrees of variability, with the exception of the highest-aptitude and below-average informants with more variable samples. In English, the average, below-average and low-aptitude test persons all had similarly higher degrees of variation than the other speakers.

German spontaneous speech was relatively uniform, as far as degrees of variability were concerned: only the average and below-average groups had less variable samples. In English, the above-average informants joined the lower-variability cluster.

The mean *F0* values in both language and speech type conditions lay within ~180-210 Hz. The greatest deviations from this trend were presented by the above-average and low-aptitude groups in German read speech (~140 Hz); below-average and native subjects in English read utterances (~140-150 Hz); above-average informants in English and the low-aptitude group in German spontaneously produced sentences (~130 Hz).

The above observations were further formally checked in the course of the ANOVA procedure. The outcome was significant for most conditions ($p<.01$), except for the German spontaneous speech ($p=.323$), which was unexpected regarding the cross-group mean *F0* mismatches. Only a few tests were significant according to the Scheffé procedure ($p<.05$), which, however, did not point out any tendencies in *d* parameter distribution of the default boundaries.

Figure 7.13 is to be referred to for the illustration parameter *d* distribution in boundary tones.

Corpus analysis by speech type and gender

Read speech

The final stage of boundary tone investigation consisted in analyzing the gender-specific group samples, split by speech type. We hereby first examined read speech in this respect, which, at the very first glance, appeared to reveal some consistent and recurring patterns in female realizations.

Full low boundary tones L-L%

This observation was most clearly supported by the distribution of L-L% boundary tones in German: both the mean *F0* values were comparable (~190-200 Hz) and the degrees of variability were almost equal in all of the groups. The corresponding productions in English, however, reflected some cross-group mismatches in female speech. More specifically, this was related to the recurrent clustering of the native, highest-aptitude and high-aptitude groups that had comparably high degrees of variability. The remaining samples were, on the other hand, characterized by an almost 50% drop in variability. There was also a separation of the mean *F0* levels: the native speakers and the two higher-aptitude groups demonstrated higher pitch in this condition than the other female speakers (~210 Hz vs. ~190 Hz, respectively).

The male samples of the L-L% targets in read speech did not allow of such easy grouping. In German, the degrees of variability were, for the most part, almost equal across the aptitude groups. However, the high-aptitude subjects had more variable *d* parameter values, coupled with a higher mean pitch (~130 Hz vs. ~100-110 Hz in the remaining groups). As opposed to this general homogeneity, the below-average male informant produced the L-L% targets in read speech with a very low degree of variability, which presented another instance of his consistently aberrant performance.

The English-language realizations allowed us to single out the high-aptitude male group performance once again – for the highest degree of variability and higher than average pitch (~140 Hz vs. ~110 Hz). The remaining groups demonstrated comparable mean *F0* values. As for the degrees of variability within these informants, we could distinguish the average and low-aptitude speakers as the groups producing more variable samples than the other male subjects, except for the outstanding high-aptitude group.

To complete the course of the empirical observations on the distribution of parameter *d* in full low boundary tones in read speech, we ran a series of ANOVAs on the data. As a result, a significant outcome in most cases was yielded ($p<.02$). When we analyzed the results of the Scheffé tests, it became evident that the L1 German speakers' performance in their native language was, for the most part, uniform. This was manifest through only sporadic statistically significant contrasts. In English, the above-average female subject with lower than average pitch and the high-aptitude male speakers with a considerably higher mean *F0,* than in the corresponding female and male groups, were separated from the latter statistically in all pertinent Scheffé contrasts ($p<.01$).

Full high boundary tones H-H%

When we embarked upon the investigation of the German-language productions of the full high boundary tones H-H%, considerable homogeneity was discovered in the female part of the read-speech corpus. Thus, most mean *F0* values were located on the same level (~230 Hz); only the low-aptitude as well the high-aptitude group productions deviated marginally (~220 Hz and ~240 Hz, respectively). The degrees of variability in *d* parameter values were also similar. The above-average subject and the highest-aptitude test persons, however, presented exceptions from this uniform pattern: the former by producing a more mean-concentrated sample; the latter – with more scattered values. The English part of the corpus in female speech allowed of a similar clustering as to the degrees of variability as the L-L% events. In particular, the native, highest-aptitude and high-aptitude informants had a similarly high degree of variation in their H-H%

targets. This was, however, replicated by the below-average group in the current experimental condition. With respect to pitch, mean *F0* values were generally higher as exhibited by the highest-aptitude and high-aptitude speakers (~240 Hz vs. ~220 Hz in the other female groups).

Several observations could be drawn from the male realizations of the H-H% boundaries. The first one was related to the performance of the below-average speaker as well as the highest-aptitude group, which was distinguished by lower degrees of variation in both languages, as compared to the remaining male subject pool.

Nevertheless the low-aptitude group's sample was identified due to its comparably low degree of variability in German and high-aptitude group realizations – in English.

Apart from their lower degree of variation, the highest-aptitude and below-average groups also demonstrated the lowest mean *F0* values in German (~120 Hz vs. ~140-180 Hz in the other groups). In English, on the contrary, these groups shared similar mean pitch with most other test persons (~120-130 Hz), except for the low-aptitude and high-aptitude informants (~150 Hz and ~170 Hz, accordingly).

All of the ANOVA tests executed on the high full boundary tone samples in read speech had significant results (p=.000). The Scheffé procedure, employed for more specific cross-group differences, only singled out the systematic differences of the low-aptitude female group as opposed to the other speakers (p<.02), except for the contrast against the above-average informants (p=.384) in German. In English, both the female and male high-aptitude group samples, characterized by higher than average *F0* values, differed significantly from most corresponding samples pertaining to the other groups (p<.02), except for the contrast with highest-aptitude group for the female speakers and the low-aptitude test persons in male speech (p>.26).

Default boundary tones %

During the investigation of default boundary tones in read speech by gender, we first analyzed female realizations of the category in question. We could therefore see that both the variability and mean *F0* patterns were, for the most part, uniform in L1 and L2. In German, only the above-average speaker departed from this trend by rendering a higher-pitch sample (with mean values of ~250 Hz vs. ~210 Hz in the other groups). The English-language part of the female read-speech corpus did not contain any major cross-group differences in *F0* levels, as the fluctuations in this regard were only marginal (~200-210 Hz in all groups). The degrees of variability were also comparable in most groups, except for the below-average and native subjects, whose *d* parameter values were more scattered around the mean.

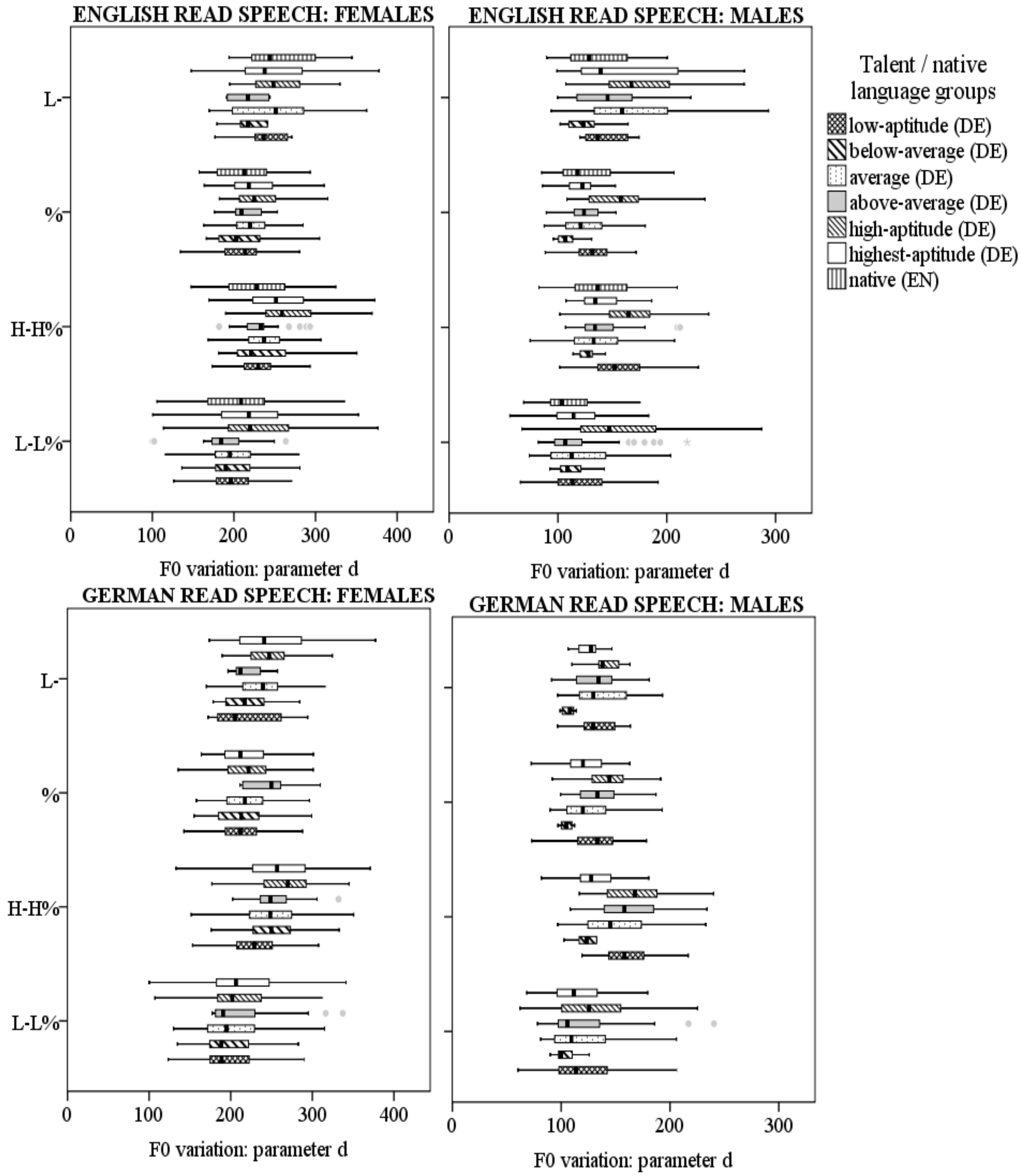


Figure 7.14: d parameter variation in boundary tones (L-L%, H-H%, %, L-) with regard to the F0 level in Hz in all the aptitude groups in English and German read speech

Male productions of the default boundary tones were less even than the female data. More specifically, the below-average subject appeared to be a clear outlier, both in English and in German: the % targets were realized on a lower pitch (~100 Hz vs. ~120-160 Hz in the remaining groups in both language conditions – the maximum values being rendered by the high-aptitude informants), as well as with a very low degree of variation. Excluding this group, the remaining male speakers realized the default boundary tones with similar degrees of variability in German. In English productions we could single out the high-aptitude and native male groups with higher degrees of variability than the remaining speakers.

The results of the ANOVA tests, which were significant in most conditions (p=.000), except for female read speech (p=.129), lent yet another insight into the data. Thus, the higher than average mean *F0* level of the above-average female speaker in German did not seem to affect the overall uniform pattern statistically (p>.05). Most cross-group contrasts in male speech were not significant either, according to the Scheffé procedure, although the high-aptitude male speakers in English consistently deviated from all the other group samples (p<.01).

Intermediate low boundary tones L-

The final boundary category, the low intermediate edge tones, did not seem to yield any clear-cut tendencies in either mean pitch levels or degrees of variability. On the contrary, male and female speakers demonstrated quite different performance in the two investigated languages.

This fact could be supported by the male data to the greatest extent: the degree of variability in English was by over 50% higher in most groups than the respective German-language samples. Looking at the male data in more detail, we could notice that the below-average, high-aptitude and highest-aptitude group samples in German all had really low degrees of variability. For the below-average subject, such idiosyncratic performance was also coupled with very low mean *F0* values (~100 Hz vs. ~130 Hz). The remaining groups had similar degrees of variability and mean pitch in German.

The English-language productions once again allowed us to single out the below-average male speaker for his idiosyncratic performance – a very low degree of variation in comparison with the other male groups, as well as low mean *F0* (~100 Hz vs. ~120-160 Hz). According to the latter parameter, i.e. mean pitch, similar behaviour was demonstrated by the native speakers, exhibiting equal values. With respect to variability, the highest-aptitude subjects yielded the most variable sample, whereby the lowest degree of variation was produced by the low-aptitude and the below-average male informants.

The German-language samples of the female subjects in German allowed of a virtual separation of the aptitude groups as to the degrees of variability: highest-aptitude and low-aptitude subjects with the highest levels; as well as a similarly lower degree of variation in the remaining groups. A different separation could be applied when we analyzed the mean *F0* values in the groups: low-aptitude, below-average and above-average speakers (~200 Hz) with lowest pitch; highest-aptitude, high-aptitude and average test persons (~220 Hz) with higher values.

A different variability / mean level picture was unveiled when investigating the L- tokens in English female speech. The average group, followed by the native and highest-aptitude subjects, had the greatest degree of variability in values in this condition. Contrary to that, the below-average and low-aptitude speakers produced the two least variable samples. Most groups were comparable with regard to the mean *F0* level (~220-230 Hz). The below-average and above-average groups, however, deviated from this general level with lower values (~200 Hz).

Despite the inter-group mismatches described above, most of the ANOVA tests did not lead to significant results ($p>.05$). Only the differences in intermediate low boundary tone tokens in German male read speech were found to be substantial ($p=.001$). Supporting these general results, no Scheffé tests were significant ($p>.40$).

Figure 7.14 is an illustration of gender-specific *d* parameter distribution in boundary tones in read speech.

At the final stage of boundary tone analysis we turned our attention to the spontaneous speech realizations of parameter *d* in all of the aptitude groups.

Spontaneous speech

Full low boundary tones L-L%

The examination of spontaneous speech started with the analysis of the most frequent category – the low full boundary tones.

In this category, the female subjects demonstrated a generally homogeneous distribution pattern in the distribution of parameter *d* values. The degrees of variability in both language conditions were similar in most groups, with a few exceptions in English and German speech. In the latter part of the experimental corpus, the above-average and average informants had slightly more variable samples. In English, the native and low-aptitude speakers produced values slightly more scattered around the mean.

The mean levels in all female groups and in both languages were also similar: marginal deviations were caused by the highest-aptitude group in German (~220 Hz vs. ~190 Hz in the remaining groups) and the above-average speaker in English (~190 Hz vs. ~200 Hz in the remaining groups).

Male samples of the low full boundaries were, in general, less variable and less consistent than the female data. Most groups demonstrated a rather low degree of variation in German, which was most common for the below-average informant. The mean *F0* concentrated around 100 Hz for the majority of speakers, although a higher level was typical of the high-aptitude group (~120 Hz). English spontaneous speech only allowed us to single out the high-aptitude test persons for their higher degree of variability; the other samples were equally less variable. The mean pitch was also shared by most groups (~100 Hz), but the high-aptitude and low-aptitude groups had higher values (~130 Hz).

The ANOVAs stated that only the female samples in English spontaneous speech were statistically similar (p=.402); the remaining conditions yielded statistically significant results (p=.000). Most specific inter-group contrasts in the Scheffé tests were not significant, which attested to the relative homogeneity in the experimental data. There were, however, two clear outliers in German, which we described above. Both the highest-aptitude female speakers and high-aptitude male subjects were identified by a higher mean pitch than the other groups. Consequently, most comparisons involving these groups lent significant outcome (p<.05), except for the contrast of above-average and highest-aptitude female groups as well as high-aptitude and below-average test persons (p>.28). The high-aptitude male group continued to produce different values in English as well (p=.000), caused by their high mean *F0*. Here, only their comparison against the low-aptitude male subjects was not significant (p=.519). The latter, in their turn, also differed from the remaining samples significantly (p<.04) due to the relatively high mean pitch values that they produced. All the other tests were not statistically significant.

Full high boundary tones H-H%

In the high full boundaries the female speakers continued to demonstrate overall consistency in variability degrees and mean *F0* levels. The general mean pitch was also shared by most groups (~230 Hz), with the exception of the high-aptitude and below-average speakers (~260 Hz). In the English-language productions, it was possible to single out the above-average speaker's idiosyncratic performance both for the low degree of variation, as well as lower than average mean pitch (~230 Hz vs. ~250 Hz). It should be noted in this regard, however, that the low-aptitude subjects had equally low mean *F0* values in their realizations. The remaining groups produced their utterances on the same mean pitch of about 250 Hz. The highest-aptitude group had the greatest variability in high full boundary tokens, which was closely followed by the native speakers with a slightly less variable sample, and, finally, the other groups.

Male productions of the H-H% contained a considerable number of cross-group peculiarities in both languages. Whereas in German the mean *F0* was similar in most groups (~140-150 Hz), with only one outlier being the above-average subjects (~190 Hz), the degrees of variability were clearly different from aptitude group to aptitude group.

The lowest degree of variability was characteristic, unsurprisingly, of the below-average subject. On the contrary, the most variable samples were found within the above-average and, especially, high-aptitude speakers.

Two variability clusters could be identified in the English spontaneous male speech: native, highest-aptitude and low-aptitude groups with lower degrees of variability and the remaining groups.

However, no clustering was possible in the mean pitch classification, as all groups had differing values from ~130 Hz (native and highest-aptitude subjects) to ~180 Hz (high-aptitude speakers).

In the course of the ANOVA procedure, it was stated that the female samples of the full high boundary tones in German were characterized by an overall agreement in statistical terms (p=.057), whereby the remaining contrasts in H-H% events were significant (p<.05). All the specific cross-group contrasts in female speech were not significant, according to the Scheffé tests (p>.31). In the male realizations, only the above-average group consistently deviated from the other samples in both languages (p<.04), except for the comparison with the below-average group in German (p=.436).

Default boundary tones %

The distribution of parameter *d* values in default boundary tones was rather even as realized by the female speakers in German. All groups had similar mean *F0* values (~200-210 Hz) and almost all produced equal degrees of variability, except for the high-aptitude and low-aptitude groups with slightly less variable samples. In the English part of the corpus, one group could be singled out as breaking the generally homogeneous pattern described for the L1 samples. Thus, the above-average subject had both very little variation in values and a slightly lower mean pitch (~190 vs. ~200-210 Hz). Notably, the native subjects had equally low mean *F0* values. Regarding the degrees of variability, the below-average group, along with the native subjects, had more variability in values than the other informants.

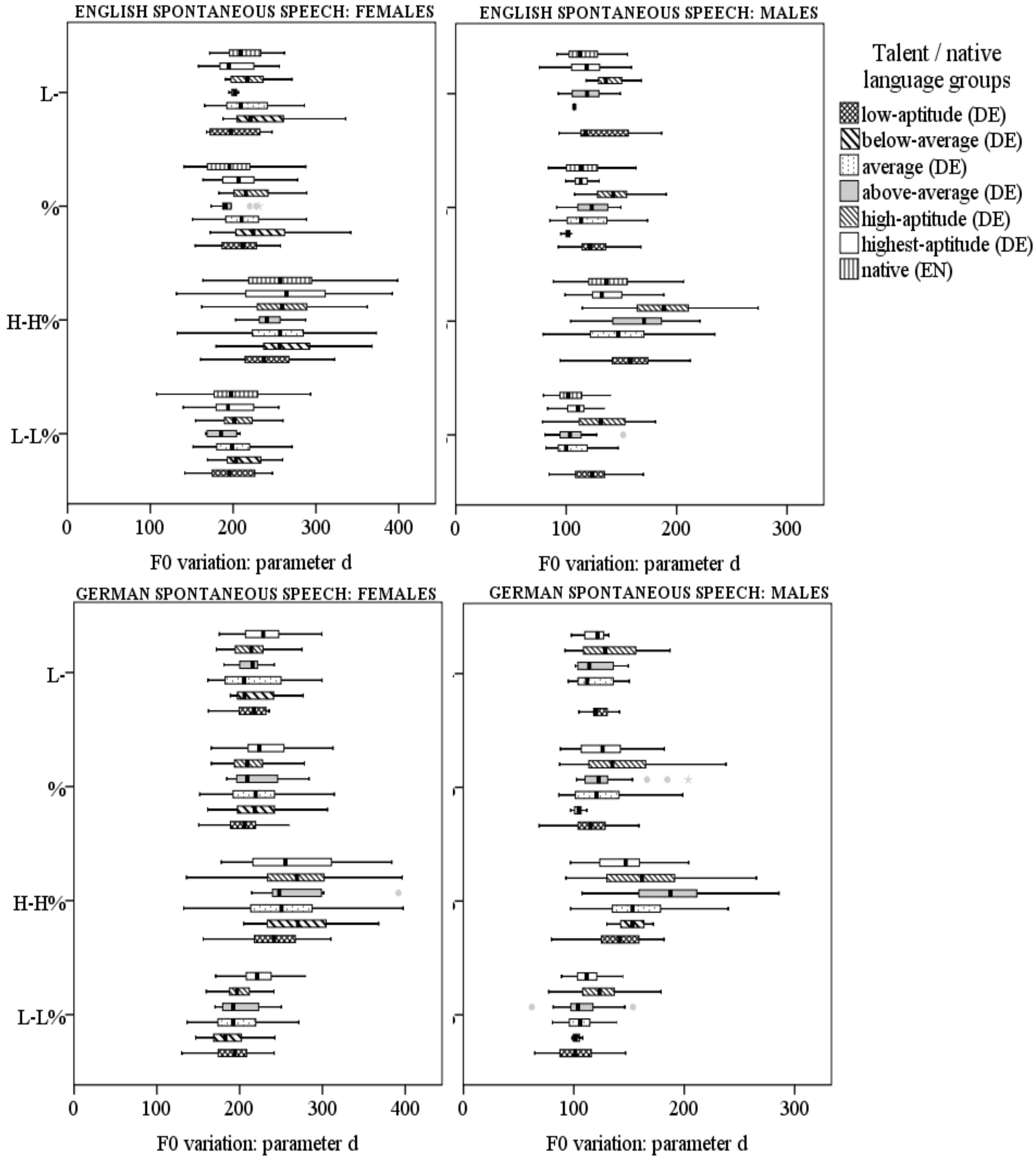


Figure 7.15: d parameter variation in boundary tones (L-L%, H-H%, %, L-) with regard to the F0 level in Hz in all the aptitude groups in English and German spontaneous speech

When analyzing the male realizations of the default boundaries register the aberrant performance of the below-average subject in both English and German utterances: there was a very little degree of variability in addition to a generally lower pitch (~100 Hz vs. ~120-140 Hz). Apart from this deviation, the higher degree of variation produced by the high-aptitude subjects was also noticeable for German. Moreover, these test persons had the highest mean pitch level in both languages (~140 Hz vs. ~130 Hz in German; ~140 Hz vs. ~120 Hz in English). The remaining pronounced peculiarity involved the highest-aptitude group's performance in English, where these subjects demonstrated almost as little variation as the below-average informant.

The default boundary tone tokens were realized in a statistically different way by all the groups in English and German, according to the respective ANOVA tests (p=.000). Similarly to read speech, only the high-aptitude male speakers appeared to be the outlying group in English spontaneous speech (p<.05 in all contrasts with the other male groups). The other comparisons were not systematic with respect to statistical significance.

Intermediate low boundary tones L-

The final category under analysis in the boundary tones were the low intermediate edge tone targets L-.

The mean *F0* levels were comparable in most groups across the two investigated languages genderwise. For the female subjects the mean pitch was registered at ~190-210 Hz, with no major deviations. In male speech, the average *F0* level was, naturally lower (~100-110 Hz). The only exception was presented by the high-aptitude informants in English spontaneous speech (~130 Hz).

One of the major observations with regard to the degrees of variability was the generally very low number of such tokens in some conditions, which was, for instance, characteristic of female above-average, male average and male below-average speakers in English spontaneous speech, as well as below-average male informant in the German spontaneous utterances. Other than this observation, the variability differences between the groups were visible, yet not systematic across the language / gender conditions.

Despite the visible mismatches between the groups in the realization of intermediate low boundaries described above, there were no statistically significant differences in that respect (p>.06). Nor did the Scheffé post-hoc tests produce any significant results (p>.13).

Figure 7.15 presents the findings in parameter *d* distribution of boundary tones in spontaneous speech.

7.2 Summary

The current chapter presents the results of an investigation into the phonetic realization of some phonological categories, namely, pitch accents and boundary tones. These categories were, in their turn, identified in the course of the phonological study described in Chapter 6, where the frequent and infrequent intonational events were singled out for the current experimental corpus. The former categories have thus been our primary focus of attention for both the phonetic and phonological levels of analysis.

We will further proceed with a summary of findings in the frequent pitch accents and boundary tones.

7.2.1 Pitch accents

With respect to the **rising L*H accents**, four PaIntE parameters appeared to be pertinent for the description – amplitude of an *F0* curve (*d*), alignment of *F0* peaks within the prominent syllables (*b*), velocity of the rising sigmoid (*a1*), amplitude of the rising sigmoid (*c1*).

The first consistent pattern discovered for the distribution of **parameter *d*** in the L*H accents was the tendency of highest-aptitude informants to realize their tokens with a greater degree of variability than the remaining groups. This observation applied to the general-corpus investigation, as well as the analysis by speech type in all experimental conditions. It should be noted in this regard that native subjects approached such performance in English read and spontaneous speech, as did the average and high-aptitude informants in their L1 and L2 read utterances. The application of the gender factor suggested that the greater degree of variability was to a larger extent the property of female subjects; native female subjects demonstrated comparable degrees of variability in English.

Still another peculiarity involving the female speakers of higher pronunciation aptitude was the fact that they realized their *d* parameter values on a higher pitch than the remaining female groups in all gender-specific conditions in both languages. Only the native female group in English read speech was close to this high pitch level. The below-average subjects, surprisingly, exhibited equally high pitch as the highest-aptitude and high-aptitude groups in spontaneous speech.

As opposed to the higher pitch values yielded by the higher-aptitude subjects, the above-average group tended for lower *F0* levels, as supported by the general-corpus and speech-type analyses. However, this trend vanished with the application of the gender factor to the data.

The final observation we could draw from the distribution of parameter d in L*H accents was related to the male below-average subject's idiosyncratic performance in male speech in general, reflected through his lower degree of variability and lower mean values. Apart from this condition, this subject also produced the least variable sample in English and German spontaneous speech.

Regarding the alignment of *F0* peaks in L*H accents, **parameter *b***, the most noticeable observation concerned the fact that the values generally appeared to be more variable in the German-language productions than in English. Moreover, the same ratio was preserved in read and spontaneous utterances taken separately, whereby the latter speech type had more variability than the former. This tendency was, for the most part, carried over to the female read realizations. The only exception here was the above-average speaker's sample with a greater degree of variability in English.

We next analyzed the velocities of the rising sigmoids within the L*H pitch accents – **parameter *a1***. As a result, no consistent patterns could be registered across the aptitude groups in all of the examined experimental conditions.

A few tendencies could be discovered during the examination of the amplitudes of rising sigmoids of L*H accents, or **parameter *c1***. The most consistent pattern hereby was the greatest degree of variability in the highest-aptitude and high-aptitude groups in the whole corpus, in read speech in general and in female read speech.

Lastly, the below-average male subject was marked with the lowest degree of variability in male speech in general, as well as German read and spontaneous utterances.

The investigation of the **H* accents** was limited in the current investigation to the description of two parameters: the amplitude of a given *F0* curve (*d*), as well as the alignment of pitch peaks within a syllable (*b*).

When examining the amplitudes of *F0* curves, **parameter *d***, we noticed some recurring phenomena with respect to the degrees of variability and the mean pitch levels.

Regarding mean *F0* values, the above-average group exhibited the lowest mean pitch of all groups in both English and German, which was closest to the native group's average pitch in English. This observation was supported during the general-corpus analysis and in read speech without regard to the gender factor. The application of the gender factor to the data allowed us to assume that female speakers were mainly responsible for this general-corpus trend: similar performance was registered in female speech in general, as well as in the corresponding read-speech realizations.

Contrary to the above tendency of the above-average subjects, high-aptitude and highest-aptitude informants tended to produce their H* targets on a higher pitch level in several experimental conditions. This was, for instance, the case with the highest-aptitude group in German speech in general.Gender-specific comparisons put forward both the highest-aptitude female speakers in

English and German and high-aptitude female group in English with the highest mean pitch levels. In male speech, only high-aptitude informants could be separated from the other subjects for the given reason in L1 and L2. The most detailed analysis of the data by speech type and gender further seconded the current trend within the higher-aptitude groups, in particular, within female speakers of high and highest pronunciation aptitude as well as the high-aptitude male subjects.

With respect to the degrees of variability, speakers of higher pronunciation aptitude, highest-aptitude and high-aptitude female subjects and high-aptitude male informants produced more variable samples in German and English speech in general, as well as in German spontaneous speech. When the gender factor was taken into account, we could see that the highest-aptitude and high-aptitude female test persons as well as the high-aptitude male subjects had the greatest degree of variability in English, similar to the native subjects' pattern. The female subjects of higher pronunciation aptitude also produced similarly variable values in German, which was approached by the average group in both languages at issue.

The gender-specific samples being further split by speech type, we could see similar variability tendencies in read and spontaneous speech. In the former speech variety, the two groups of higher pronunciation aptitude once again demonstrated greatest degrees of variability, similarly to the native group in English and average speakers in German. However, no such distinction could be observed for spontaneous speech. In male realizations, we could cluster the highest-aptitude and native subjects in English read speech due to the fact that their samples were more variable than those in the remaining groups.

A final note should be made about the below-average male subject's performance, as it deviated from the other group samples quite observably in several experimental conditions: in German male speech in general, as well as the corresponding read and spontaneous realizations, the degree of variability was visibly lower than in the remaining male groups.

Turning our attention to **parameter *b***, we did not notice many consistent trends across the groups with respect to the degrees of variability. Only in gender-specific examinations could we see that the below-average female group was marked by an increased degree of variation than the remaining female subject pool in read and spontaneous utterances, as well as female speech in general. Among male speakers, the low-aptitude group was, on the other hand, characterized by the smallest degree of variability in all comparisons involving the gender factor.

The final general observation regarding the distribution of *b* parameter values in H* accents concerned a greater degree of variability in the German-language data in all L1 German groups in all experimental conditions, with the exception of the below-average male speaker in spontaneous speech.

The **rising-falling tones L*HL** revealed some recurrent pattern with regard to both *F0* level and degrees of variability in some of the aptitude groups in **parameter *d***. More specifically, the low-aptitude informants exhibited reduced variability and lower mean pitch in both languages at issue during the general-corpus analysis and in female speech not split by speech type.

Contrary to the low *F0* level displayed by low-aptitude subjects, native speakers realized their prominent syllables on the highest pitch. This peculiarity was caused by female realizations to the greatest extent. However, this difference of native speakers was, for the most part, not statistically significant.

The only consistent pattern in male speech in L*HL accents was related to the below-average subject's performance in both English and German, where he produced the lowest mean *F0* values of all groups.

Looking into the distribution of ***b*** **parameter** values in L*HL accents, we could only see a few cross-group peculiarities in the gender-specific analyses. The general trend for male and female speakers with respect to the alignment of *F0* peaks was the fact that L1 German subjects exhibited more variability in their mother tongue than in English. Low-aptitude male subjects, however, departed from this pattern by producing a less variable sample in German. In general, this group was singled out due to lower degrees of variability in both languages. The only other subjects deviating from the otherwise consistent pattern were above-average speakers in English with a more variable sample than the rest of the female subject pool.

Regarding **parameter *c2***, one of the peculiarities regarded the greater degree of variability within native, highest-aptitude and high-aptitude informants in general, as well as the corresponding female segments of these aptitude groups. Female native and highest-aptitude test persons also realized their rising-falling targets on a higher mean pitch in English than the remaining groups.

With respect to the **simple low targets L***, in **parameter *d*** we discovered lower mean *F0* in the above-average group in English and German in the whole corpus, which was not, however, supported by gender-specific analysis. The only instance of conspicuous performance in the data divided by gender was presented by the high-aptitude male group with both a higher degree of variability and higher mean pitch than the other groups.

The distribution / frequency pattern for the low targets was also relatively uniform in **parameter *b***. This was reflected in comparable degrees of variability in most groups and language conditions. In contrast to such homogeneity, above-average speakers in the general-corpus analysis, as well as the male segment of this group, exhibited more variation in English. The given group also displayed lower mean values in their L2 speech, without regard to the gender factor. The same was true for the low-aptitude group in German.

The **falling H*L accents** revealed a similar trend as the low targets described above with respect to **parameter *d***: we could register a lower mean *F0* level in L1 and L2 speech of the above-average group in the general corpus compared with the remaining groups, which in English was also characteristic of the native subjects. Such behaviour of above-average speakers did not find proof in gender-specific samples.

With the application of the gender factor, however, we could notice a greater degree of variability in the two higher-aptitude groups in both English and German, which was also replicated in the native speaker performance for H*L accents. The high-aptitude and highest-aptitude groups also realized their falling-tone targets on a higher mean pitch in English. In German, only the highest-aptitude group was distinguished due to their highest mean pitch. In male speech, only high-aptitude informants demonstrated higher values in both languages.

We could not register any clear trend for the alignment of *F0* peaks, parameter b, within H*L accents.

The main differences in **parameter *c2*** were found during the gender-specific analysis. More specifically, in female speech in English the highest-aptitude, high-aptitude and native speakers could once again be clustered due to their higher degrees of variability and higher pitch than the other groups. Unexpectedly, the low-aptitude female group displayed similarly high *F0* values.

In male speech, only the below-average speaker could be singled out due to his generally lower degree of variability.

7.2.2 Boundary tones

The investigation of boundary tones was limited to the description of **parameter *d*** distribution, since the application of other PaIntE parameters in this respect appeared to be problematic.

For the most frequent boundary tone category, the **low full boundary tones L-L%,** we could reiterate the observation about the above-average group's lower mean *F0* values during the general-corpus analysis, as well as in spontaneous speech samples without regard to gender. Notably, native speakers almost replicated this low mean pitch in English speech in the respective conditions.

To the gender-specific peculiarities belonged the male high-aptitude group performance due to greater degrees of variability and higher mean *F0* than those demonstrated by the remaining groups. The only exception to this trend was the German spontaneous speech condition, where the high-aptitude male informants were only singled out for their high pitch, but not a more variable sample.

One more observation during the gender-specific analysis concerned the native, highest-aptitude and high-aptitude female informants in the English part of the corpus regardless of speech type, as well as in English read speech: these samples marked by greater variability and higher mean *F0* values.

The analysis of the **full high boundaries H-H%** allowed us to single out the above-average group with their lower than average mean *F0* during the general-corpus and speech-type investigations. However, as with the L-L% edge tones, this trend of the above-average group was not congruent with gender specification.

With the gender factor applied, we could notice the higher degree of variability within the highest-aptitude informants in female speech in general, as well as in the respective read-speech samples. In the English part of the corpus, the native female test persons produced similarly variable values. Apart from this peculiarity, the highest-aptitude group, as well as the high-aptitude female speakers, demonstrated significantly higher mean pitch than the other female groups in English female productions as a whole and in read speech.

The final group that could be singled out due to their realization of L-L% boundary tones were the male speakers of high pronunciation aptitude, who had higher than average pitch in German and English male samples without the account of the speech type factor, as well as in their L2 and L1 read speech.

When we turned our attention to the **default boundary category %**, the above-average group was once again singled out for their lower pitch during both the general-corpus and speech type analysis in both the investigated languages.

In gender-specific analyses, the below-average male speaker exhibited recurrently aberrant behavior, involving lower pitch and less variable values in parameter ***d*** than the other male speakers.

Completing the analysis of boundary tones, we examined the distribution of **L- intermediate boundary tones**, within which no significant trends were discovered.

Chapter 8

Theoretical, methodological and experimental aspects of intonational variation and pronunciation aptitude

The current dissertation is an investigation of intonational variation and its relation to pronunciation aptitude. In the course of this investigation, we first attempted to lay the theoretical ground for the subsequent experimental endeavours. Thus, Chapters 1-3 served as a conceptual foundation for the given study, and Chapter 4 allowed us to select its methodological paradigm. Based on these premises, we undertook an experimental study described in Chapters 6-7. This chapter, in its turn, is supposed to bring the theoretical and methodological assumptions in line with the results of the actual experiment.

We will therefore discuss several aspects of the key phenomena of intonational variation and pronunciation aptitude: theoretical, methodological and experimental. Within the first aspect, we will look into its major perspectives, i.e. general linguistic and SLA. Thus, we will explore the reflection of the theories discussed in Chapters 1-3 in the experimental results. Further on, the methodological aspect will be examined as to the suitability of its underlying premises for the current study – classification of subjects into aptitude groups, the data annotation approach, as well as the PaIntE model for *F0* parametrization. Finally, the issues raised specifically in the given dissertation will be addressed. We will, first of all, confront the hypotheses put forth in Chapter 5, in order to see if they find any support in our data. The discussion will be finalized by summarizing any additional implications following the results of the current dissertation.

8.1 Theoretical aspects

8.1.1 General linguistic perspective

In Section 2.1 we reviewed five major linguistic theories and their treatment of intonational variation. More specifically, we investigated the underlying concepts of Structural Phonology, Generative Linguistics, Optimality Theory, Prototype Theory and Exemplar Theory.

As a result, a conclusion was drawn that Exemplar Theory (e.g. Nosofsky 1986, Pierrehumbert 2001, Bybee 2005, Wade et al. 2010) can predict intonational variation most accurately, due to the fact that it takes into account concrete speech representations, which, in turn, allows of realistic descriptions of intonational categories. The experimental study described in Chapters 6-7 does not appear to contradict with this initial assumption. Focusing on the realization of frequent categories, we could describe their functioning through the sets of their constituent exemplars.

Although infrequent classes were not investigated with similar thoroughness, we would assume that Exemplar Theory might prove to be an equally good or even better predictor of mechanisms governing infrequent categories as well. One of the main challenges about infrequent categories is their instability: it is often difficult to apply boundaries or ranges, within which exemplars can be freely produced, due to the presence of a substantial number of outliers. Therefore, conventional statistical methods used for frequent categories are not quite suitable for infrequent events. Thus, the investigation of the simple low L* targets as well as intermediate low L- and high H- boundary tones, registered as infrequent in the course of the current investigation, did not result in any clear distribution / realization patterns, either on the phonetic or the phonological level, when the same procedure as for frequent categories was employed. Examining exemplars of infrequent classes directly might account for the variability within them more accurately, thereby revealing some inherent tendencies in the data.

Still another perspective into the infrequent categories could be seen if we analyze the behaviour of the six investigated aptitude groups, especially with the application of the gender factor. The bigger groups (frequent human categories) exhibited relatively stable variability and frequency of occurrence patterns throughout the course of the study: for example, highest-aptitude and high-aptitude female subjects rendered similar degrees of variability in most investigated phonological categories and PaIntE parameters; average subjects of both genders consistently remained within the phonetic and phonological variability limits set by most other groups. Contrary to that, one-person groups (infrequent human categories) often deviated from the overall trends in an inconsistent way.

More specifically, the degree of **phonetic variability** within the above-average female and below-average male subjects could be smaller, greater as well as equal to that in the other group samples. Evidently, the number of tokens produced by the above-average female and the below-average male informants was smaller than in the remaining groups, which would predict an unalterably lower degree of variability. This was, in fact, the case with the male below-average male speaker in a number of experimental conditions, including but not limited to the following:

- with regard to parameter *d* of L*H accents in English and German spontaneous speech;
- in parameter *c1* of L*H accents of all male productions in English and German, as well as the corresponding German read and spontaneous utterances;
- during the examination of parameter *d* of H* accents in German male speech, with and without regard to the speech type factor;

- in the realization of parameter *d* of the default % boundary tones;
- in parameter *d* values of the low intermediate boundary tones L- in both language conditions, without account of the speech type factor and in read utterances.

The above-average female subject, on the other hand, did not differ substantially from the remaining female groups with regard to degrees of variability in the realization of most categories. She did, however, produce a more variable sample in parameter *b* of the rising-falling L*HL pitch accents in English.

A question arises in this regard whether what we observed within those single-person groups was real or "false" variability. An example of the latter would be an infrequent category, where individual tokens are located at a considerable distance from one another, but their range might coincide or even exceed that in a frequent category, with a denser population of tokens. This fact might explain visibly equal or greater degrees of variability demonstrated by the below-average male and above-average female speakers in some conditions.

It appears that the investigation of such categories in terms of individual exemplars could shed more light into the functioning of infrequent language events, since their description does present certain difficulties.

Regarding the above considerations on the nature of phonetic variability in infrequent categories, we cannot but recall the theoretical arguments mentioned in Chapter 2, put forth in previous studies. The logical conclusion to draw from several of such investigations is the fact that frequent categories should be characterized by a greater degree of variability (e.g. Pierrehumbert 2001; Schweitzer and Möbius 2004), since frequent categories accumulate a larger number of exemplars during their development than infrequent ones. Wade et al. (2010), contrarily, predicted an opposite trend on the same phenomenon stating that frequent categories are more stable than the infrequent ones and are therefore characterized by a smaller degree of variation.

Following the implications of the current study, we would assume that it is not only the number of exemplars accumulated, it is also their range of distribution and dispersion that affects the outcoming ratios between the degrees of variability in frequent and infrequent phonological events. Thus, an infrequent category sample may equal or even exceed a frequent phonological event in variability, should the tokens in the former category be more scattered.

Looking at the **phonological level**, we could also find a few instances of idiosyncratic behaviour within the single-person groups, which was especially characteristic of the below-average male test person. First of all, he employed the rising tone category L*H least extensively of all female and male L1 German groups, except for the native subjects in English with a similarly low number of L*H events. Such performance was especially non-compliant, regarding the generally high percentage of rising tones in German. Reminiscent of this observation was the below-average male subject's realization of the high full boundary tones, where he was also singled out for a lower number of H-H% tokens than the other groups. On the contrary, with respect to the simple high targets H*, the below-average male speaker deviated from a generally homogeneous distribution pattern by producing a larger number of the events in question, which reached ~85% in English spontaneous utterances, exceeding the corresponding percentage within the native informants, for whom this category was more common overall than for the L1 German subjects.

All of the above examples point out to the instability of single-person groups (infrequent human categories), which may, on the one hand, comply with the general sample or depart from it, on the other. This fact, in turn, is also related to the notion of individual variation and the heterogeneity of human populations, where some individuals would be perfect representatives of general trends, while others would render outlying instances. Further on, such outliers may, with category evolution, become regular category members. It appears in this regard that Exemplar Theory would be most suitable for the description of infrequent categories and their development, as it is meant to analyze individual speech representations.

Once a frequent category is established, in all its aspects, i.e. phonetic, phonological, social and so on, its members populate a general common space. Inside this space, one can establish boundaries, within which exemplars can be freely produced. That would then agree with the description of a category in Structural Phonology (e.g. Saussure 1916; Trubetzkoy 1958; Bloomfield 1933). However, the latter theory would still fail to categorize any exemplars emerging beyond the category boundaries during its development. The structural approach thereby appears more suitable for synchronic descriptions of frequent categories, while Exemplar Theory takes account of the diachronic aspect as well.

With respect to the remaining theories we reviewed in Section 2.1, we did not find any additional evidence as to their pertinence in the explanation of intonational variation.

The general linguistic aspect was, undoubtedly, one of the cardinal points in the current study. Nevertheless due to the fact that we analyzed productions of the six aptitude groups in two languages, English and German, the SLA perspective appears to be of equal importance for the interpretation of experimental results.

8.1.2 SLA perspective

Section 2.2 presented an overview of the major second language acquisition theories in view of their implications towards cross-linguistic variation; more specifically, the degrees of variability in non-native speech. The outcome of the current study did, in fact, support several of these theoretical paradigms in at least some of their underlying ideas.

Basic Variety Theory

We could therefore relate some of the experimental results to the Basic Variety Theory (Klein & Perdue 1997). According to BV, learners develop an error-free reduced system of a given L2, suitable for their communicative needs. From a developmental perspective, SLA is characterized by a gradual increase in the degree of speech variability as the learner's proficiency reaches higher levels: learners with a low proficiency level should demonstrate a low degree of variation, whereas the speech of highly proficient individuals should be highly variable, approaching the target-language variation levels, i.e. native-speaker performance.

The current investigation but partially confirms this idea. Thus, we could apply the theory's implications on the phonetic level – the phonological investigation appeared to follow a different logic.

We could indeed see greater degree of phonetic variability demonstrated by the highest-aptitude and high-aptitude female groups as compared to the other L1 German female speakers. Such a high degree of variation was also closest to the native female group sample, which was evident in several experimental conditions. As an example, we can mention the following:

- *c1* parameter values of L*H accents analyzed by speech type and gender;
- parameter *d* analysis of H* events by gender and in female read speech in both languages;
- parameter *d* samples of L*HL accents split by gender in English and German;
- parameter *d* investigation by gender in the two languages in H*L accents;
- parameter *c2* in the English-language female samples of H*L accents;
- gender-specific analysis of L-L% boundary tones in the general English-language corpus and in English read speech.

Those mentioned above are not the only conditions exhibiting the clustering of the highest-aptitude, high-aptitude and native female groups. Sections 7.1-7.2 provide a detailed account of all findings on the phonetic level.

Such proximity of higher-aptitude and native speakers was perhaps the sole consistent evidence in favour of the Theory of Basic Variety. The other results lead to contrary conclusions. Thus, although we could group female subjects with respect to the degrees of variability, this pattern was not supported by male speech. We could indeed often observe greater degrees of variability, coupled with higher mean pitch, within the high-aptitude male informants, but their performance was rather outstanding.

Another counterargument to BV, logically emerging from the results of the experiment, is the absence of any gradual increase in the degrees of variability with the rising level of pronunciation aptitude. Within the female subjects, to whom the theory in question is only applicable, only the highest-aptitude and high-aptitude groups were characterized by consistently more variable samples. No other recurring cross-group differences could be discovered for the remaining speakers, despite the fact that the theory would predict the lowest degrees of variability for the low-aptitude and below-average groups. However, such instances were only random.

Other theories, on the other hand, could explain the absence of gradualism across the aptitude groups. The results of the current study therefore resonated to the greatest extent with the Interlanguage Theory.

Interlanguage

Selinker's Interlanguage Theory (Selinker 1972) and its major concepts provide a relatively solid explanation to SLA processes.

One of the most pertinent concepts for the current study is the effect of *transfer*. We could indeed register instances of consistent phonological transfer from L1 into L2, which was most characteristic of the events ending on high pitch – rising tones L*H and the high full boundary tones H-H%. This appeared to be one of the cardinal tendencies of the current investigation due to the greater distribution of the events at issue in German (e.g. Mayer 1995) and their much sparser employment in English.

Most L1 German groups thus transferred their native-language **phonological** distribution pattern, i.e. a higher percentage of the rising tones and full high boundary tones, into English. This tendency was, however, suppressed rather consistently by subjects of highest and high pronunciation aptitude. They did comply with the general distribution pattern in German; in their L2 productions, on the other hand, the highest-aptitude and high-aptitude speakers reduced the number of rising L*H accents and H-H% boundary tones. This, in turn, appeared to approach the two higher-aptitude groups to the native speakers. To be more precise, the female subjects were responsible for this trend in the corpus: the observation in question was corroborated by gender-specific analyses of L*H and H-H% events, as well as by female read speech with regard to the rising tones. Male productions did not support the given tendency in a consistent manner.

Summarizing, we observed a transfer of a typical category by most female groups into L2, resistance to the transfer phenomenon within the highest-aptitude and high-aptitude female speakers as well as their accommodation to the L2 frequency of occurrence pattern.

Logically following the above observation, we could say that most L1 German groups *overused* typical native-language category, L*H accents and H-H% boundary tones, and *avoided* the production of events more characteristic of English – simple H* targets and low full boundary tones L-L%. Indeed, the distance in percentages of the latter events (simple high targets and low full boundary tones) between the native male speakers, on the one hand, and the other male groups, on the other, was quite pronounced. In female speech, the highest-aptitude, high-aptitude and native subjects were respectively separated from the remaining subject pool with their higher numbers H* and L-L% tokens.

Whether individual **phonetic** realizations of pitch accents underwent the processes of transfer, overuse or avoidance remains unknown. We would assume that most groups would also make no significant distinction between English and German intonation, but this idea needs further perceptual verification.

One other concept is worth mentioning with respect to the Interlanguage Theory – that of *systematic* and *nonsystematic variation*. The former type of variability has been explored in a number of studies (Couper-Kuhlen 2001; Couper-Kuhlen & Selting 1996; Local 1989, 1992; Local, Kelly & Wells 1989, Lowry 2002, Grabe 2004, Halliday 1967, Ladd 1980). Nonsystematic variants remained beyond the scope of analysis, for the most part. Our investigation touched upon both aspects of variability. We would hereby relate systematic variation to the phonological level in those cases, where the choice of an intonational event affects the semantic content of a given utterance. The semantic aspect was not analyzed within the current study, but it does contain potential for further research. The phonetic variants would, as we would assume intuitively, adhere to nonsystematic phenomena, prone to random interchangeability. Nevertheless it would be an additional experimental challenge to test, whether free intonational variants actually do not have any influence on the semantic domain.

The current dissertation contains detailed descriptions of both systematic (ToBI accent and boundary tone distribution) and nonsystematic variants (the phonetic realization of ToBI accents and boundary tones within the six PaIntE parameters). However, the above-mentioned problematic aspects, alongside with any other potential issues, can find further theoretical as well as experimental development.

Some explanation for free variation is found in the foundations of the Diffusion Model.

Diffusion Model

The main idea behind the Diffusion Model (Gatbonton 1978) was the fact that newly acquired L2 forms coexist with L1 exemplars in free variation, and later they are sorted into their appropriate contexts. We cannot disprove this general postulation with the experimental results, as far as the phonetic level of analysis is concerned. As mentioned previously in the interpretation of the Theory of Basic Variety, we could only separate the highest-aptitude, high-aptitude and native female subjects with consistently higher degrees of variability. No steady variability tendencies could be registered for the remaining groups, which might be accounted for by the fact that both properly L1 and L2 tokens populate overlapping categories in the English and German speech of our test persons. Consequently, that results in comparable degrees of variability between the groups. The highest-aptitude and high-aptitude female speakers, contrary to that, seem to have overcome the "sorting out" of the L1 tokens and have continued enriching their L2 proper exemplar inventory.

Perceptual Assimilation Model (PAM)

Some explanation of the SLA processes can be traced back to the Perceptual Assimilation Model (Best 1995). PAM posits, in particular, that an L2 token is susceptible to merging into an L1 category, the smaller the distance between such a pair of phonological events is. This assumption appears to fit quite well within the current framework, since English and German have often been reported as typologically close languages with respect to intonation (e.g. Kuhlmann 1952, Schubiger 1965, Esser 1978, Scuffil 1982, Grabe 2000). This is also supported by the proximity of the two ToBI paradigms.[45]

For the current investigation results, we might assume that the L*H and H-H% categories might have been assimilated by most groups across L1 and L2, due to their active consistent transfer into English. A perceptual investigation is needed, however, in order to state whether such assimilation resulted in a greater number of exemplars, which can be classified as an L1 or and L2 token.

[45] See Section 5.2.1 for a description of the German and English transcription systems within the ToBI framework.

Second Language Speech Learning Model (SLM)

Second Language Speech Learning Theory (Flege 1995) also sheds some light onto the SLA processes at play in the current investigation. It does, therefore, explain the reason for L1 transfer into L2. As mentioned earlier, English and German are close typologically. Such proximity, in Flege's opinion, hampers category discrimination. Consequently, the undiscriminated L1 / L2 categories undergo transfer. For instance, the rising tones L*H and high full boundary tones H-H% are common for both English and German, for which reason they are normally not discriminated by L1 German subjects, so their native-language tokens replace the proper English ones.

Markedness Differential Hypothesis and Fundamental Difference Hypothesis

The only two theories, which we reviewed in Section 2.2 and which do not contain any direct links with the results of the current investigation are the Markedness differential Hypothesis (Eckman 1977) and the Fundamental Difference Hypothesis (Bley-Vroman 1989). The former is more related to the semantic domain, and the latter denies the very existence of high pronunciation aptitude in adults.

Nevertheless, in general, we can state that most SLA theories account for this or that actual SLA process. According to the analysis provided above, the Interlanguage Theory is perhaps the greatest predictor of L2 behaviour by non-native speakers, although the other theories explain important cross-language aspects as well.

We have dwelled upon the theoretical premises of the current investigation and their reflection in the actual experimental outcome. The suitability of the methodological basis will be discussed in Section 8.2 below.

8.2 Methodological aspects

Several methodological aspects are of crucial important for the current study: the classification of speakers according to their degrees of pronunciation aptitude, the phonological data labelling procedure as well as the *F0* parametrization model.

8.2.1 Classification of subjects into pronunciation aptitude groups

The data pertaining to part of the subjects recruited for the current investigation has been used in our previous research (Anufryk 2009; Anufryk & Dogil 2009; Anufryk 2008; Anufryk et al. 2008), where we applied a slightly different speaker classification, as opposed to the current study. Previously, we used three-group analysis, i.e. the test persons were split into three main groups based on Matthias Jilka's assessment of their pronunciation aptitude and proficiency (see Jilka 2009; Jilka et al. 2007) – the below-average, average and the above-average subjects. However, our intuition was that these groups were not homogeneous.

Therefore, the subjects were subdivided into six groups, applying a finer gradation, to account for the heterogeneity and individual variability to a greater extent. The resulting classification included highest-aptitude, high-aptitude, above-average, average, below-average and low-aptitude groups.

The challenge for gender-specific analysis was the fact that two groups contained only one subject. We thus observed one above-average female and one below-average male speaker. One might argue that results of single-person groups may not be representative of the whole population of such speakers. We would not, in any way, contend this fact. We did, however, decide upon retaining the samples of these two groups within the analysis, as they might be examples of individual variation – the heterogeneity of speech.

This assumption was corroborated by the experimental results: whereas the above-average female subject, for the most part, complied with the general trends, the male below-average speaker deviated from them on both the phonetic and phonological levels of analysis (see examples in Sections 6.3, 7.2 and 8.2).

The more detailed classification of subjects also allowed us to see the closest proximity of the highest-aptitude and high-aptitude female groups, demonstrating similar performance in most conditions, which was, on the one hand, closest to the behaviour of the native subjects and, on the other hand, made those speakers distinct from the other female groups. Unfortunately, we could not register any such regularity in male speech.

8.2.2 Annotation approach

Section 5.2.1 described the data annotation approach in the current investigation. Generally, we adhered to the ToBI framework (e.g. Silverman et al. 1992; Pitrelli et al. 1994; Mayer 1995; Grabe 2001), due to its solid theoretical basis in the autosegmental-metrical paradigm (e.g. Liberman 1967; Bruce 1977; Pierrehumbert 1980; Gussenhoven 1983; Ladd 1998) and the fact that it has been applied to a substantial amount of modern natural languages, including German and English. The challenge in a cross-linguistic was in the interaction between the systems. Thus, both English and German were described separately, regardless of the cross-linguistic implications. An additional complication was the existence of British (Grabe 2001) and American ToBI (Silverman et al. 1992; Pitrelli et al. 1994) as well several autosegmental-metrical models for German (e.g. Wunderlich 1988; Féry 1988, 1993; Uhmann 1988, 1991; Grice & Benzmüller 1994; Mayer 1995).

We decided to choose the Stuttgart model for German (Mayer 1995), based on the pre-analysis of the experimental data, which could be more easily classified within this model. Both the English and American notations were also applied to the labelling of the corpus. The reasoning behind this approach was the fact that we were investigating non-native speech, which, on the one hand, contained instances of L1 transfer and, on the other hand, it was often difficult to classify it within the English-language varieties. We thus adhered to a mixed inventory of pitch accents and boundary tones, which was supposed to allow us to label the non-native realizations realistically and see whether the native speakers' realizations were all uniform or whether they could be classified within a different accent / language variety.

Indeed, with respect to the latter observation, we could single out some tokens in the native realizations as pertaining to a typically German category of the default % boundary tone – unchanging pitch at the end of a phrase. One might argue that those were representatives of the American ToBI H-L% category, but these events presuppose a lowering of pitch to approximately the mid level, whereas the German default boundary is just unchanging pitch at IP boundaries. It was this latter case that we could observe in native speech.

Overall, the annotation approach allowed of a relatively realistic labelling of experimental tokens for their further phonetic and phonological analyses.

8.2.3 *F0* parametrization model

As a result of the review of the major prosodic models in Chapter 4, we decided to carry out the analyses in the current investigation within the PaIntE Model (Möhler 1998, 2001). There were two main reasons behind this choice. Firstly, this model complies with the annotation ToBI scheme. Secondly, it renders detailed phonetic information about each separate *F0* curve in terms of six parameters: *a1* and *a2* (the velocity of the rising and the falling sigmoids, respectively); *b* (the alignment of the *F0* peak within a three-syllable window); *c1* and *c2* (the amplitudes of the rising and the falling sigmoids, accordingly); *d* (the amplitude of the *F0* excursion within an individual curve).

The PaIntE Model appeared to render detailed information about each *F0* curve identified on the phonological level. The six PaIntE parameters therefore exposed some systematic cross-group trends, which was especially characteristic of the frequent ToBI events.

8.3 Experimental aspects

8.3.1 Validation of hypotheses

The final stage of experimental vs. theoretical investigation consisted in the validation of hypotheses laid down in Chapter 3, which were, in turn, based on the theoretical premises of Chapters 1-2. The results of the current investigation corroborated some of these assumptions, while discarding the others. We will analyze each of these hypotheses separately below.

Hypothesis 1: *L2 speech of less proficient (lower than average level of pronunciation aptitude) speakers should exhibit a smaller degree of variability in a foreign language.*

As seen from the summaries in Sections 7.2 and 8.2, no consistent phonetic variability patterns could be discovered for most groups, except for the highest-aptitude and high-aptitude female speakers, who consistently demonstrated higher degrees of variability (e.g. in *c1* parameter values of L*H accents analyzed by speech type and gender; parameter *d* analysis of H* events by gender and in female read speech in both languages; parameter *d* samples of L*HL accents split by gender in English and German; parameter *d* investigation by gender in the two languages in H*L accents; parameter *c2* in the English-language female samples of H*L accents; gender-specific analysis of L-L% boundary tones in the general English-language corpus and in English read speech). Consequently, degrees of variability in the remaining groups in the above conditions appeared to be lower, thus validating the first hypothesis.

Hypothesis 2: *In an opposite fashion, advanced-proficiency (higher than average level of pronunciation aptitude) individuals should be more variable in their speech due to their higher acuity in mastering L2 exemplars and employing them in communication.*

The current hypothesis was supported by the female realizations pertaining to the highest-aptitude and high-aptitude groups in their realization of phonetic PaIntE parameters in a number of experimental conditions (see examples mentioned for Hypothesis 1). In Section 8.2 we also linked such an increase in degrees of variability to the underlying concepts of the Basic Variety Theory.

Hypothesis 3: *Average-proficiency (and average-aptitude) speakers are expected to exhibit average variability levels, relative to the higher degrees of variation within the more talented speakers and less variable samples of the less talented individuals.*

In fact, the average subjects, especially female, exhibited variability degrees close to those of the highest-aptitude and high-aptitude groups in several conditions in the phonetic-level investigation:

- parameter *d* of L*H accents: by speech type in both language conditions, in the speech of average informants in the whole experimental corpus in L1 and L2 and in English female read speech;
- parameter *d* of H* accents: in female speech in general in English and German as well as in female read-speech realizations in German;
- parameter *d* of low full boundaries H-H%: in English in the whole corpus and in L1 and L2 read speech;
- parameter *d* of the default boundary tones %: in both languages without regard to the factors of speech type and gender and in English and German read speech;
- parameter *d* of the intermediate low boundaries: in female read speech in English.

On the whole, we could state the relative stability within the average speakers as to the degrees of phonetic variability in that this group did not present any consistently deviating performance. Vice versa, their values fit within the general patterns in most experimental conditions.

Hypothesis 4: *Degrees of variation between speakers of average and higher than average pronunciation aptitude are unpredictable and can only be verified through the analysis of real-life speech data.*

In fact, as seen from the analysis of the SLA theories in Section 8.2, only the highest-aptitude and high-aptitude female subjects were singled out for their greater degrees of variability in the realization of PaIntE parameters in conditions specified. The variability ratio between the remaining groups was not distinct and their performance was not consistent: the corresponding samples could exhibit both higher and lower degrees of variation in the investigated pitch accent and boundary tone categories, with and without the application of the two influence factors.

Hypothesis 5: *On a proximity scale, speakers of higher than average aptitude and proficiency should approach most closely the native-speakers' variation pattern. Average as well as lower than average levels of aptitude and proficiency should correlate with lower degrees of variability, establishing a greater distance of such speakers from the proper L2 performance.*

The final experimental hypothesis found evidence in female speech. Indeed, highest-aptitude and high-aptitude female informants displayed higher degrees of phonetic variability than the other L1 German female groups, which was also closest to the native subjects in performance on the phonetic level in a number of conditions mentioned previously: in *c1* parameter values of L*H accents analyzed by speech type and gender; parameter *d* analysis of H* events by gender and in female read speech in both languages; parameter *d* samples of L*HL accents split by gender in English and German; parameter *d* investigation by gender in the two languages in H*L accents; parameter *c2* in the English-language female samples of H*L accents; gender-specific analysis of L-L% boundary tones in the general English-language corpus and in English read speech.

To summarize the results of the hypotheses validation, we ought to point out that the experimental evidence presented above only pertained to the phonetic level. The phonological analysis, on the other hand, did not appear to follow these assumptions. We will therefore discuss the implications of the phonological investigation, alongside with any additional phonetic issues, in Subsection 8.3.2.

8.3.2 Additional phonological and phonetic implications

8.3.2.1 Phonological level

Distribution of ToBI pitch accents and boundary tones

The analysis of the ToBI pitch accents and boundary tones allowed us to relate the degrees of variability within the aptitude groups to the L1 / L2 accommodation vs. transfer phenomena. Thus, there was no straightforward correlation of a greater degree of variability with a higher level of pronunciation aptitude. We could rather observe the replication, or non-replication, of the native-language distribution patterns within the informants.

As a result, we found that most L1 German groups tended to transfer their L2 behaviour in English, whereas speakers of highest and high pronunciation aptitude attempted to accommodate to the L2 model, thereby approaching the native subjects. This tendency was especially prominent in the L*H pitch accent category during the general corpus investigation, where the highest-aptitude and high-aptitude informants reduced the percentages of this typical German pitch accent in their L2 speech, as compared to German. This made their performance closest to the native subjects, exhibiting even lower numbers of rising tones. The remaining groups, on the contrary, simply carried over their L1 frequency of occurrence pattern into English. The same trend of the highest-aptitude and high-aptitude, alongside with native, subjects, on the one hand,

and all the remaining German speakers, on the other, was further supported by read-speech data without regard to gender and in female read and spontaneous speech.

The high full boundary tones H-H% exposed a similar pattern as the L*H accents. Here, highest-aptitude and high-aptitude informants were once again clustered with native speakers due to their lower percentages of H-H% events in English, as opposed to the other L1 German groups. This tendency was observed in the whole corpus, in read speech in general, in female speech without account of the speech type factor and in female read-speech realizations.

The higher-aptitude speakers therefore accommodated to the English-language model by reducing the number of rising tones L*H and full high boundaries H-H%, as these categories enjoy a scarcer distribution in English. An opposite trend was discovered for the low full boundary tones - a category more frequent in English speech. The highest-aptitude and high-aptitude female speakers, vice versa, increased the percentage of L-L% to fit the native-speaker pattern, during the general gender-specific analysis and in read-speech samples. The remaining female and male groups did not modify their behavior across the languages.

Summarizing the results of pitch accent and boundary tone distribution in the given experimental corpus, we would like to conclude that phonological variability most closely relates to the cross-language frequency of occurrence ratios. Thus, if two languages include identical categories in their phonological inventories, the frequency of occurrence and distribution of such events might be different, and the speakers ought to follow these patterns. A seen in non-native speech, however, only speakers of higher than average pronunciation aptitude (in the given study – most consistently female subjects) appeared to accommodate to the L2 model observed within native speakers, whereby the subjects with a lower level of pronunciation aptitude fail to discriminate between their L1 and L2 in this respect.

As evidenced from the experimental results, the phonetic and phonological levels follow a different logic with regard to the degrees of variability in individual pitch accents and boundary tones. The phonetic results relate to the quantitative aspect of variability – its greater / smaller degrees; the phonological analysis rather reflects the interlanguage transfer and accommodation phenomena in intonation. When we investigated an even more global level, i.e. intonational text interpretation, we could not apply any of those earlier observations to the corresponding analysis. Instead, the data appeared to suggest still another pattern across the aptitude groups.

Global prosodic text interpretation

The results of the global phonological investigation on text level, i.e. the comparison of ToBI transcriptions of the two texts under analysis for the uniformity / variability of prosodic text interpretation (see Section 6.2 and Subsection 6.3.2) did not seem to suggest any trends similar to the previous levels of analysis. Thus, the first tendency was the lower degree of cross-speaker variability in all German-language transcriptions of L1 German speakers as compared to greater variability in all their English productions. One of the logical explanations to this fact could be the general comfort of the subjects with their mother tongue and their subsequent neutral interpretation of the experimental texts, which also suggest some common text interpretation patterns in a given community of speakers

The English versions of the two investigated texts allowed us to discover some other consistent patterns of prosodic text interpretation across the groups. Thus, the native and average subjects demonstrated the highest degrees of cross-speaker agreement (expressed by lower Levenshtein values), regardless of the gender factor in the "North Wind and the Sun" samples as well as in female speech of "Mrs. McWilliams". One might possibly explain such results by these subjects' relatively neutral attitude to the task, which yielded neutral realization templates. The highest-aptitude and low-aptitude speakers, on the other hand, demonstrated the greatest degrees of variability in prosodic text interpretation, which was especially evident in their productions of "Mrs. McWilliams". We would assume different explanations for this fact for the two groups. The highest-aptitude informants might have strived for a more varied and "interesting" interpretation of the experimental texts, which entailed using varying intonational means and resulted in a higher degree of cross-speaker disagreement. The low-aptitude speakers, on the other hand, regarding their lower proficiency level, might not have been certain, which event to employ in this or that context, hence the higher Levenshtein distance values attesting to greater disaccord within the group in this respect.

The underlying reasons for greater / smaller degrees of cross-speaker variability in prosodic text interpretation across the groups does not appear to be guided by any phonetic or phonological factors per se. Rather, the results reflect the speaker's extra-linguistic predisposition towards the experimental task (neutral – for average and native subjects; more involved – for the highest-aptitude group) as well as their proficiency level in English. Thus, the inventory of language units at a subject's disposal, coupled with his or her knowledge of their application in an appropriate context, defines the degrees of variability. Lower-aptitude subjects might therefore apply prosodic events inappropriately and differently from speaker to speaker, which creates a greater degree of cross-speaker variability.

8.3.2.2 Phonetic level

An additional tendency to the ones discussed with respect to the experimental hypotheses as well as the general linguistic and SLA theories was the fact that female speakers of highest and high pronunciation aptitude realized their PaIntE parameters on a higher mean *F0* level, by ~20-30 Hz, than the other groups in a number of experimental conditions (see Section 7.2.1 for a summary). This was, for instance, the case with the realization of PaIntE parameter *d* of L*H accents in all gender-specific comparisons. The same trend was also supported by the distribution of parameter *d* in H* accents and H-H% boundary tones in English female speech.

This finding could be related to the above tendency of the highest-aptitude group for a higher degree of variability in prosodic text interpretation. It appears that the subjects not only attempted to make their speech more variable phonologically, they also raised their mean pitch, higher than native speakers, which is also commonly related to a more lively speech style. Several earlier studies also referred to such a phenomenon of overhitting the L2 target, although on the segmental level (e.g. Bohn & Flege, 1992).

8.3.3 Comparison of the current results with our earlier research

As mentioned earlier in Chapter 5, we conducted several investigations of the data rendered by our subjects prior to committing to the current study. After comparing the results of these investigations, we could state the presence of recurrent trends in the data.

Thus, it was, first of all, related to the tendency of highest-aptitude and high-aptitude female subjects to suppress their native-language distribution pattern of L*H accents when speaking English, thereby approaching native speaker's performance. This finding was confirmed in all our earlier studies (Anufryk 2008; Anufryk et al. 2008; Anufryk 2009; Anufryk & Dogil 2009), where speakers of higher than average pronunciation aptitude accommodated to the proper English distribution model to the greatest extent. However, these investigations did take into account cross-gender peculiarities in the distribution of pitch accents. Apart from that, only read speech was analyzed. The current study, on the other hand, applied a more sophisticated classification of subjects by levels of pronunciation aptitude, coupled with an in-depth analysis of male and female speech, both read and spontaneous.

The second consistent trend we encountered regarded the higher overall mean pitch of parameter *d* within the highest-aptitude and high-aptitude female subjects. We found support for this observation in two previous comprehensive studies of read speech. The first one of them only made a general statement about higher pitch of more talented speakers in general, without account of the gender factor, in L*H accents (Anufryk & Dogil 2009). The other investigation (Anufryk 2009) examined the data gender-specifically: male speakers of higher pronunciation aptitude were found realize their H* tokens on a higher mean pitch, whereby the same was true of the female subjects in this respect. The given study found more evidence for this trend in female speech. More specifically, parameter *d* had higher mean *F0* values in L*H accents, H* accents and H-H% boundary tones, as produced by highest-aptitude and high-aptitude female subjects (see Section 7.2.1 for a summary).

Finally, to the recurring phonetic trends also belonged the peculiarity of highest-aptitude, high-aptitude and native female subjects to render more variable samples of some PaIntE than the other groups. Our earlier research singled out this tendency for parameter *d* both for more talented speakers in general (Anufryk & Dogil 2009) for L*H accents and genderwise (Anufryk 2009) for L*H and H*L events. In the current dissertation, the given trend could be confirmed for parameter *d* and expanded to several other parameters. To summarize, the following conditions corroborated the observation on greater variability within higher-aptitude speakers: *c1* parameter values of L*H accents by speech type and gender; parameter *d* of H* events by gender and in female read speech in both languages; parameter *d* of L*HL accent by gender in English and German; parameter *d* by gender in the two languages in H*L accents; parameter *c2* in the English-language female samples of H*L accents; gender-specific analysis of L-L% boundary tones in the general English-language corpus and in English read speech.

We could also support the findings in the general prosodic text interpretation, when looking at one of the earlier studies (Anufryk 2009). In that investigation, speakers of less than average and higher than average pronunciation aptitude demonstrated a greater degree of disagreement in the prosodic text interpretation of "The North Wind and the Sun", as compared with average and native speakers. It should be noted in this regard, however, that the number of participants in that earlier investigation was lower than in the current study. As discussed in Section 6.3.2, we discovered similar patterns both within the "North Wind and the Sun" and the abstract from "MrsMcWilliams".

CONCLUSION

The current dissertation is a cross-discipline, linguistic and SLA, study of pronunciation aptitude and its relation to intonational variation. In the course of the study, we first laid down the theoretical and methodological foundations for the subsequent experimental investigation and proposed several hypotheses following our interpretation of the major linguistic and second language acquisition theories. Further on, the results of this investigation were presented, both on the phonetic and phonological levels, for subjects of varying pronunciation aptitude.

Linguistically, the focus of our attention lay within the frequent phonological categories, and these appeared to accord with some general linguistic concepts about category architecture, mostly of the Exemplar Theory and Structural Phonology. We concluded that the former is best-suited for a dynamic description of category functioning and formation, as it takes account of individual phonetic representations. The structural phonological approach is capable of category definition statically: any outlying instances cannot be explained.

Infrequent phonological events remained beyond the scope of the study. It appears, however, that infrequent classes provide insight into the diachronic perspective, i.e. category development, alongside with rendering synchronic information on how they function at a certain point in time. It would therefore be interesting to dedicate attention to infrequent categories in our future work.

From the SLA perspective, the results of the current experimental investigation fell within the predictions of several major second language acquisition theories – Basic Variety, Interlanguage, Diffusion Model, Perceptual Assimilation Model and Second Language Speech Learning Theory. Our study, however, could contribute to the understanding of their underlying principles in a number of ways.

First of all, we could see that pronunciation aptitude and proficiency is not linearly related to variation, as set forth in the Basic Variety Theory. Our data only implied the difference of more phonetically capable, and proficient, subjects with regard to variability. More talented / proficient female subjects indeed realized some categories with a greater degree of phonetic variation than the less talented / proficient individuals. We could not draw any distinctions between the latter as to the different degrees of variability.

There remain several unsolved issues in this regard. Thus, it would be interesting to verify the current results within a larger sample of speakers to see whether only talented subjects consistently exhibit outstanding performance, reflected in higher phonetic variability in intonation. Apart from that, since only female subjects recurrently followed the trend for higher variability, male subject data should be examined in greater detail to prove or disprove the absence of this trend in their speech.

We could also apply the concepts of the Interlanguage Theory to the current investigation. The phenomena of transfer, avoidance and overuse found their evidence in our phonological results. Additionally, our study contributed to the exploration of non-systematic variability on the phonetic and phonological levels, as both of them allow of free interchangeability of intonation tokens. The interrelation between systematic and non-systematic aspects of intonational variation, however, still remains unexplored. It is commonly posited that intonation has its semantics, though not as strictly defined as in the lexical domain. At the same time, intonational categories can be in free variation. Phonetic variants, on the other hand, should be predominantly freely interchangeable, although it is not clear whether phonetic properties have any influence on semantics. All these questions call for further development.

The **methodology** of the current study relied on several important elements: subject classification, data annotation approach and intonation modeling procedure. All three elements appeared to be suitable for the investigation undertaken.

As compared to our earlier research, we applied a more sophisticated classification of subjects into aptitude groups (six instead of three), which allowed of a more detailed examination of data, accounting for individual variation. Indeed, the preservation of single-person groups enabled us to see that some speakers comply with the general trends (the above-average female subject), while others demonstrated outlying performance (below-average male informant). Larger aptitude groups appeared to be generally more stable in behavior than those single speakers.

With respect to data annotation, the eclectic approach of combining the German and English ToBI pitch accent and boundary tone inventories had the benefit of seeing the SLA transfer phenomena in L1 German groups, when they would employ a typically German category in an L2 context. It also showed some tendencies in the speech of native speakers. The most prominent observation in this respect was their rather common use of the traditionally German default boundary tone category.

For the data analysis, we employed the PaIntE Model (Möhler 1998, 2001), which created six-dimensional representations of each individual *F0* curve carrying a boundary tone or pitch accent. Although we mainly explored parameter *d* (amplitude of an *F0* curve), other parameters have potential for exploration as well. In particular, an interesting finding regarded parameter *b*, which had more variability within the German groups in both English and German, whereby the samples in the latter language were visibly more variable. This observation, as well the examination of the remaining PaIntE parameters can become the subject of a separate further study.

Several **hypotheses** laid down in Chapter 3 found experimental evidence. We could thus state that the speech of more proficient and talented individuals is characterized by a greater degree of phonetic variability than that produced by less talented and proficient speakers. This greater degree of phonetic variability also brings more capable individuals closer to the native speakers in performance. The given hypotheses, however, only appeared to govern the phonetic level. The results of the phonological investigation imply a closer link with the SLA processes of transfer and accommodation. The greatest degree of such accommodation was found within higher-aptitude subjects.

Summarizing the implications of the various aspects of the current study, we can suggest its **application** in several domains.

Firstly, the findings of the current investigation contribute to the exploration of the general linguistic phenomenon of intonational variation on both phonetic and phonological levels.

Second of all, understanding a learner's performance is crucial for teaching purposes. Having examined speakers of varying pronunciation aptitude, we also provided an insight into their performance with respect to intonation. This information could be used for optimizing and individualizing the teaching process, which should, in turn, unveil an individual's potential to the greatest extent.

We are also convinced that the results of the current study can be used in speech modelling and technology. Within the first application, one may, for example, model a speaker of a certain level of pronunciation aptitude and proficiency, where the intonational aspect can become the primary focus as well as one of the elements. In that latter case, one might speak of creating models of artificial intelligence covering various language and other skills. The findings of the current study may also be applied in cross-language speech recognition. It is commonly known that accented speech is more problematic for such models. Knowing the mechanisms behind it might make the creation of hybrid models, capable of better recognition, possible.

LIST OF ABBREVIATIONS

BV: Basic Variety Theory (Klein & Perdue 1997), a theory describing variability in second language speech, described in detail in Section 2.1.1.

CANAL-F: Cognitive Ability for Novelty in Acquisition of Language Test (Grigorenko, Sternberg & Ehrman 2000), a modern language aptitude test, described in Section 1.1.

DLAB: the Defense Language Aptitude Battery (Petersen & Al-Haik 1976), a military modern language aptitude test, described in Section 1.1.

F0: fundamental frequency, an acoustic parameter, the lowest frequency of a periodic waveform.

INTSINT: International Transcription System for Intonation (Hirst & Di Cristo 1999; Hirst et al. 2000), a system for labeling intonation events, described in Subsection 4.2.3.

IPO: Input-Process-Output Model ('t Hart & Cohen 1973; 't Hart & Collier 1975; 't Hart et al. 1990), an *F0* approximation approach, described in Subsection 4.2.4.

L1: a common abbreviation for a speaker's first, i.e. native, language.

L2: a common abbreviation for an individual's second, i.e. foreign, language.

LOR: length of residence, one of the factors of L2 exposure.

MLAT: Modern Language Aptitude Test (Carroll 1981), one of the first traditional aptitude tests, described in Section 1.1.

OT: Optimality Theory (e.g. Prince & Smolensky 1993; McCarthy & Prince 1993), one of the general linguistic theories, described in Section 2.1.

PaIntE: Parametric Intonation Events Model (Möhler 1998, 2001), a paradigm adopted for intonational analysis in the current study, described in Subsection 4.3.2.

PAM: Perceptual Assimilation Model (Best 1995), a model of phonetic acquisition through the prism of perception, described in Subsection 2.3.3.

PLAB: Pimsleur's Language Aptitude Battery (Pimsleur 1966), a language aptitude test, described in Section 1.1.

RFC: Rise / Fall / Connection Model (Taylor 1995), a phonetic model for the description of intonation, described in Subsection 4.3.2.

RMSE: root-mean-square error, a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing being modeled or estimated.

SLA: second language acquisition.

SLM: Second Language Speech Learning Theory (Flege 1995), an SLA theory, explaining the difficulties of cross-language learning, described in Subsection 2.3.3.

SPE: 'The Sound Pattern of English' (Chomsky & Halle 1968), a seminal generative linguistic work by Chomsky and Halle, mentioned in Subsection 2.1.2.

ToBI: Tones and Break Indices (e.g. Silverman et al. 1992; Pitrelli et al. 1994; Mayer 1995; Grice & Benzmüller 1994; Baumann et al. 2000), an autosegmental-metrical framework for the description and labeling of intonation, an approach employed in the current investigation, described in Subsection 5.2.1.

LIST OF TABLES

LIST OF FIGURES

REFERENCES

Abrahamsson, N. & Hyltenstam, K. (2008), 'The Robustness of Aptitude Effects in Near-Native Second Language Acquisition', *Studies in Second Language Acquisition* **30**, 481-509.

Abu-Rabia, S. & Kehat, S. (2004), 'The Critical Period for Second Language Pronunciation: Is there such a thing? Ten case studies of late starters who attained a native-like Hebrew accent', *Educational Psychology* **24**, 77-98.

Au, S. Y. (1988), 'A critical appraisal of Gardner's social-psychological theory of second-language (L2) learning', *Language Learning* **38**, 75-100.

Beckman, M. E. & Pierrehumbert, J. (2000), 'Positions, probabilities, and levels of categorization', 'Keynote address, Eighth Australian International Conference on Speech Science and Technology', Canberra.

Best, C. T. (1995), A Direct Realist View of Cross-Language Speech Perception 'Speech Perception and Linguistic Experience: Issues in Cross-Language Research', Timonium, York Press.

Bialystok, E. (2002), 'On the reliability of robustness: A reply to DeKeyser', *Studies in Second Language Acquisition* **24**, 481-488.

Birdsong, D. (2005), 'Nativelikeness and non-nativelikeness in L2A research', *IRAL* **43**, 319-328.

Birdsong, D. (1992), 'Ultimate attainment in second language acquisition', *Language* **68**, 706-755.

Black, A. & Hunt, A. (1996), *F0* contours from ToBI labels using linear regression, *in* ''Proceedings of International Conference on Spoken Language Processing, Philadelphia', 1385-1388.

Bley-Vroman, R. Susan M. Gass & J. Schachter, ed., (1989), *Linguistic Perspectives on Second Language Acquisition*, Cambridge University Press, chapter What is the logical problem of foreign language learning?, 40-68.

Bloomfield, L. (1933), *Language*, New York: Holt.

Bod, R. (2006), 'Exemplar-Based Syntax: How to Get Productivity from Examples', *The Linguistic Review* **23**.

Boersma, P. (2007), 'Cue constraints and their interactions in phonological perception and production', Rutgers Optimality Archive 944.

Boersma, P. (1998), *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives*, Holland Academic Graphics.

Boersma, P. (1997), How we learn variation, optionality, and probability, *in* 'IFA Proceedings 21', 43-58.

Bohn, O.-S. (1995), *Speech Perception and Linguistic Experience: Issues in Cross-Language ResearchSpeech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, MD, York Press, chapter Cross-Language Speech Perception in Adults: First Language Transfer Doesn't Tell It All, 279-304.

Bolinger, D. Greenberg, J. H., ed., (1978), *Universals of Human Language. Vol. 2, Phonology*, Stanford: Stanford University Press, chapter Intonation Across Languages, 471-524.

Bolinger, D. (1958), 'A Theory of Pitch Accent in English', *Word* **14**, 109-149.

Bolinger, D. (1951), 'Intonation: Levels Versus Configurations', *Word* **7**, 199-210.

Bongaerts, T. (1999), In Birdsong, D. (ed.) *Second language acquisition and the critical period hypothesis*, Mahwah, NJ: Erlbaum, chapter Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners, 133-159.

Bongaerts, T., Mennen, S. & van der Slik, F. (2000), 'Authenticity of Pronunciation in Naturalistic Second Language Acquisition: The Case of Very Advanced Late Learners of Dutch as a Second Language', *Studia Linguistica* **54**, 298-308.

Burling, R.Winitz, H., ed., (1981), *Native Language and Foreign Language Acquisition*, New York: New York Academy of Sciences, chapter Social constraints on adult language learning, 279-290.

Burstall, C. (1975), 'Factors affecting foreign language learning: A consideration of some recent research findings', *Language Teaching Abstracts* **8**, 5-25.

Bybee, J. (2006), 'From usage to grammar: The mind's response to repetition', *Language* **82**(4), 711-733.

Carroll, J. B.Glaser, R., ed., (1962), *Training research and education*, Pittsburgh: Univ. of Pittsburgh Press, chapter The prediction of success in intensive foreign language training.

Carroll, J. B.Walsh, T. & Diller, K., ed., (1981), *Individual differences and universals in language learning aptitude*, London: Newbury House Publishers, chapter Twenty-five Years of Research on Foreign Language Aptitude, 83-118.

Carroll, J. B. (1953), *The study of language: A survey of linguistics and related disciplines in America*, Cambridge, Mass.: Harvard Univ. Press.

Celce-Murcia, M. (1991), Teaching English as a second or foreign language'Teaching pronunciation', Heile & Heile Publishers, 136-153.

Chomsky, N. (1965), *Aspects of the Theory of Syntax*, Cambridge: MIT Press.

Chomsky, N. & Halle, M. (1968), *The Sound Pattern of English*, Harper & Row, New York, chapter A Sketch of English Phonology and Phonological Theory, 15-55.

Chomsky, N. & Miller, G. (1963), 'Introduction to the Formal Analysis of Natural Languages', *Handbook of Mathematical Psychology*, 271-321.

Clement, R. & Krudenier, G. (1985), 'Aptitude, Atittude and Motivation in Second Language Proficiency: A Test of Clement's Model', *Journal of Language and Social Psychology* **4**, 21-37.

Cochrane, R. (1977), 'The Acquisition of /r/ and /l/ by Japanese Children and Adults Learning English as a Second Language', PhD thesis, University of Connecticut.

Colletta, S.P., Clement, R. & Edwards, H.P. (1983), *Community and parental influence: effects on student motivation and French second language proficiency*, Quebec : International Center for Research on Bilingualism.

Couper-Kuhlen, E. (2001), 'Interactional prosody: High onsets in reason-for-the-call turns', *Language in Society* **30**, 29-53.

Couper-Kuhlen, E. & Selting, M. (1996), *Prosody in conversation: Interactional Studies*, Cambridge University Press.

Crompton, A. (1982), *Slips of the tongue and language production*, Berlin: Mouton, chapter Syllables and segments in speech production, 663-716.

Crookes, G. & Schmidt, R. (1989), 'Motivation: Reopening the research agenda', Technical report, University of Hawaii Working Papers in ESL 8.

Cruz-Ferreira, M. (1999), 'Prosodic Mixes', *International Journal of Bilingualismnull* **3, 1**, 1-21.

Cruz-Ferreira, M. (1989), 'A test for non-native comprehension of intonation in English', *International Review of Applied Linguistics in Language Teaching* **27(1)**, 23-39.

Crystal, D. (1969), *Prosodic Systems and Intonation in English*, Cambridge University Press.

De Bot, K. (1983), 'Visual feedback of intonation: effectiveness and induced practice behavior', *Language and Speech* **26**, 331-350.

De Bot, K. (1980), 'Evaluation of intonation acquisition: A comparison of methods', *International Journal of Psycholinguistics* **7**, 81-92.

De Bot, K. & Mailfert, K. (1982), 'The Teaching of Intonation: Fundamental Research and Classroom Applications', *TESOL Quarterly* **16(1)**, 71-77.

De Bot, K., Smith, E. K. &. S., ed., (1986), *Crosslinguistic Influences in Second Language Acquisition*, Oxford: Pergamon, chapter The Transfer of Intonation and the Missing Data Base, 110-119.

De Pijper, J. R. (1983), *Modelling British English intonation*, Dordrecht: Foris.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998), 'Evidence in favour of a broad framework for pronunciation instruction', *Language Learning* **48(3)**, 393-410.

Dickerson, L. & Dickerson, W. Corder & Roulet, ed., (1977), *The Notions of Simplification, Interlanguages and Pidgins*, Neuchatel: Faculte des Lettres, chapter Interlanguage Phonology: Current Research and Future Directions.

van Dommelen, W. A. & Husby, O. (2009), 'The perception of Norwegian word tones by second language speakers', *Journal of the Acoustical Society of America,* **125-4**, 2773-2773.

Dornyei, Z. (2005), *The Psychology of the Language Learner. Individual Differences in Second Language Acquisition*, Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Eckman, F. (1977), 'Markedness and the Contrastive Analysis Hypothesis', *Language Learning* **27-2**, 315-330.

Eckman, F., Elreyes, A. & Iverson, G. (2003), 'Some principles of second-language phonology', *Second Language Research* **19**, 169-208.

Ehrman, M. (1998), 'The Modern Language Aptitude Test for Predicting Learning Success and Advising Students', *Applied Language Learning* **9**, 31-70.

Elliot, A. (1995), 'Foreign Language Phonology: Field Independence, Attitude, and Successof Formal Instruction in Spanish Pronunciation', *Modern Language Journal* **79**, 530-542.

Ellis, R. (1985), 'Sources of Variability in Interlanguage', *Applied Linguistics* **6, 2**, 118 - 131.

Fathman, A. (1975), 'The relationship between age and second language productive ability', *Language Learning* **25**, 245-253.

Felix, S. F. (1985), 'More evidence on competitive cognitive systems', *Second Language Research* **1**, 47-72.

Flege, J. (1988), 'Factors affecting degree of perceived foreign accent in English sentences', *Journal of the Acoustical Society of America* **84**, 70-79.

Flege, J. (1987), 'A critical period for learning to pronounce foreign languages?', *Applied Linguistics* **8**, 162-177.

Flege, J. & Liu, S. (2000), 'Reevaluating the effect of length of residence (LOR) on adults' performance in a second language', *Journal of the Acoustical Society of America* **107-5**, 2803-2803.

Flege, J. E. & Strange, W., ed., (1995), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, MD, York Press, chapter Second Language Learning Theory, Findings, and Problems, 233-277.

Flege, J., Bohn, O-S., & Jang, S. (1997), 'The effect of experience on nonnative subjects' production and perception of English vowels', *Journal of Phonetics* **25**, 437-470.

Flege, J.E. & Bohn, O.-S. (1989), 'The perception of English vowels by native Spanish speakers', *Journal of the Acoustical Society of America* **85**.

Flege, J.E. & Hillenbrand, J. (1984), 'Limits on phonetic accuracy in foreign language speech production', *Journal of the Acoustical Society of America* **76-3**, 706-721.

Flege, J.E. & Hammond, R. (1982), 'Mimicry of Non-distinctive Phonetic Differences between Language Varieties', *Studies in Second Language Acquisition* **5-1**, 1-17.

Flynn, S. & Manuel, S. Eubank, L., ed., (1991), *Point Counterpoint: Universal Grammar in the Second Language*, John Benjamins Publishing Co, chapter Age Dependent Effects in Adult L2 Learning: An Examination of Critical Period Hypothesis, 117-146.

Fujisaki, H. (1983), Dynamic Characteristics of Voice Fundamental Frequency in Speech and Singing'The Production of Speech', New York: Springer-Verlag, 39-55.

Féry, C. (1993), *German intonational patterns*, Tuebingen: Niemeyer.

Féry, C. (1988), Rhythmische und tonale Struktur der deutschen Intonationsphrase, 'Intonationsforschungen', Tuebingen: Niemeyer, 41-64.

Féry, C. & Truckenbrodt, H. (2005), 'Sisterhood and tonal scaling', *Studia Linguistica* **59**, 223-243.

Gardner, R. C. (1985), *Social Psychology and second language learning: The role of attitudes and motivation*, London: Arnold.

Gatbonton, E. Rutherford, W., ed., (1978), *Second Language Acquisition and Second Language Universals*, Amsterdam: John Benjamins, chapter Patterned phonetic variability in second language speech: a gradual diffusion model.

Goldinger, S. (1998), 'Echoes of echoes? An episodic theory of lexical access', *Psychological Review* **105**, 251-279.

Goldsmith, J. (2001), An overview of autosegmental phonology, *in* Charles W. Kreidler, ed., 'Phonology. Critical Concepts', Routledge, London & New York, 382-425.

Goldsmith, J. (1990), *Autosegmental and metrical phonology*, Cornwall: TJ Press.

Goldsmith, J. (1976), 'Autosegmental phonology', PhD thesis, MIT.

Goldsmith, J. & Laks, B. (to appear), 'Generative Phonology: its origins, its principles, and its successors', To appear in The Cambridge History of Linguisticsedited by Linda Waugh and John E. Joseph.

Grabe, E., Gilles, P. & Peters, J., ed., (2004), *Regional Variation in Intonation.*, Tuebingen: Niemeyer, chapter Intonational variation in urban dialects of English spoken in the British Isles, 9-31.

Grabe, E. (2001), 'The IViE Labelling Guide'.

Granstrom, B. (1997), Applications of intonation: an overview, *in* 'Proceedings of ESCA workshop on intonation: Theory, models and applications', 21-24.

Grice, M. (1995), 'Leading tones and downstep in English', *Phonology* **12**, 183-233.

Grice, M., Baumann, S. & Benzmüller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In: Jun, Sun-Ah (ed.) Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press.

Grice, M. & Benzmueller, R. (1994), 'Transcription of German using ToBI-tones - the Saarbruecken System', Technical report, University of Saarbruecken.

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. (2000), 'A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test', *Modern Language Journal* **84**, 390-405.

Grossberg, S. (2003), 'Resonant neural dynamics of speech perception.'(CAS/CNS-TR-02-008), Technical report, Department of Cognitive and Neural Systems and Center for Adaptive Systems Boston University.

Guiora, A., Beit-Hallahmi, B., Brannon, R., Dull, C. & Scovel, T. (1972), 'The effects of experiementally induced changes in ego states on pronunciation ability in a second language: An exploratory study', *Comprehensive Psychiatry* **13**, 421-427.

Gussenhoven, C. (1983), *On the Grammar and Semantics of Sentence Accents*, Indiana: IULC, chapter A semantic analysis of the nuclear tones of English.

Halliday, M. A. K. (1967), *Intonation and grammar in British English*, The Hague: Mouton.

Hayes, B. (to appear), 'Phonological Acquisition in Optimality Theory: The Early Stages', To appear in Kager, Rene, Pater, Joe, and Zonneveld, Wim, (eds.), Fixing Priorities: Constraints in Phonological Acquisition. Cambridge University Press.

Hermann, G. (1980), 'Attitudes and success in children's learning of English as a second language: the motivational vs the resultative hypothesis', *English Language Teaching Journal* **34**, 247-254.

Hirschberg, J. & Pierrehumbert, J. (1986), Intonational Structuring of Discourse, *in* 'Proceedings of the 24th Meeting of the Association for Computational Linguistics', 136-144.

Hirst, D. & Di Cristo, A., ed. (1999), *Intonation Systems: A Survey of Twenty Languages*, Cambridge University Press.

Hirst, D.J., Di Cristo, A. & Espresser, R. Horne, M., ed., (2000), *Prosody : Theory and Experiment*, Dordrecht, Pays-Bas: Kluwer, chapter Levels of description and levels of representation in the analysis of intonation, 51-87.

Hjelmslev, L. (1974), *Prolegomena zu einer Sprachtheorie*, Max Hueber Verlag, München.

House, D. (1997), Temporal alignment categories of accent-lending rises and falls, *in* 'Proceedings of Eurospeech, Rhodos', 879-882.

Hyltenstam, K. (1978), 'Variability in Interlanguage Syntax', Technical report, Department of General Linguistics, Lund University, Sweden.

Hyltenstam, K. (1977), 'Implicational patterns in interlanguage syntax variation', *Language Learning* **27-2**, 383-411.

Ioup, G., Boustagui, E., El Tigi, M. & Moselle, M. (1994), 'Reexamining the critical period hypothesis. A case study of successful adult SLA in naturalistic environment', *Studies in Second Language Acquisition* **16**, 73-98.

Ioup, G. & Tansomboon, A. & Weinebrger, S. H., ed., (1987), *Interlanguage Phonology the Acquisition of Second Language Sound System*, Cambridge, MA, chapter The Acquistion of Tone: A Maturational Perspective.

Jakobson, R. & Halle, M. (1971), *Fundamentals of Language*, The Hague.

Jenner, B. R. (1976), 'Interlanguage and Foreign Accent', *Interlanguage Studies Bulletin*, 166-195.

Johnson, J.S. & Newport, E. (1989), 'Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language', *Cognitive Psychology* **21**, 60-99.

Jun, S.-A., ed. (2006), *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford: Oxford University Press.

Kingdon, R. (1973), *The Groundwork of English Intonation*, London: Longman.

Klein, W. & Perdue, C. (1997), 'The Basic Variety (or: Couldn't natural languages be much simpler?)', *Second Language Research* **13**, 302-346.

Kohler, K. (1995), 'The Kiel Intonation Model (KIM), its implementation in TTS synthesis and its applications to the study of spontaneous speech', Technical report, Kiel University.

Kohler, K. (1977), *Einführung in die Phonetik des Deutschen*, Berlin: Erich Schmidt Verlag.

Krashen, S. (1982), *Principles and Practice in Second Language Acquisition*, Pergamon.

Krashen, S.D., Long, M.H., & Scarcella, R. (1979), 'Age, rate, and eventual attainment in second second language language acquisition', *TESOL Quarterly* **13(9)**, 573-582.

Kuhl, P. (1991), 'Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not', *Perception & Psychophysics* **50**, 93-107.

Ladd, D. R. (1980), *The structure of intonational meaning*, Bloomington: Indiana University Press.

Ladd, R. (1998), *Intonational Phonology*, Cambridge University Press.

Lado, R. (1957), *Linguistics Across Cultures*, University of Michigan Press.

Lehiste, I. (1970), *Suprasegmentals*, Cambridge: MIT Press.

Lenneberg, E. Fodor, J. & Katz, J., ed., (1964), *The Structure of Language*, NJ: Prentice Hall., chapter The Capacity of Language Acquisition.

Lenneberg, E. (1967), *Biological foundations of language*, New York: Wiley & Sons.

Lepetit, D. (1989), 'Cross-Linguistic Influence in Intonation: French / Japanese and French / English', *Language Learning* **39-3**, 397-419.

Levelt, J. (1989), *Speaking: From intonation to articulation*, Cambridge: MIT Press.

Levenshtein, V. (1965), 'Binary codes capable of correcting deletions, insertions, and reversals', *Doklady Akademii Nauk SSSR* **163**, 845-848.

Liberman, M. & Pierrehumbert, J. (1984), Intonational invariance under changes in pitch range and length, *in* Mark & Richard T. Oehrle Aronoff, ed., 'Language, sound, structure: studies in phonology presented to Morris Halle by his teacher and students', Cambridge: MIT Press.

Liberman, P. (1967), *Intonation, Perception and Language*, Cambridge, Mass.: MIT Press.

Local, J., Auer, P. & di Luzio, A., ed., (1992), *The contextualization of language*, Amsterdam: John Benjamins, chapter Continuing and restarting, 272-296.

Local, J. (1989), *Doing Phonology: observing, recording, interpreting*, Manchester: Manchester University Press.

Local, J., Kelly, J. & Wells, W. H. G. (1986), 'Towards a phonology of conversation: Turn-taking in Tuneside English', *Journal of Linguistics* **22-2**, 411-437.

Long, M. H. (1990), 'Maturational constraints on language development', *Studies in Second Language Acquisition* **12**, 251-285.

Lowry, O. (2002), 'The stylistic variation of nuclear patterns in Belfast English', *Journal of the International Phonetic Association* **32-1**, 33-42.

Marinova-Todd, S. H. (2005), 'Proficiency in the Home and School Languages of Bilingual Children''XIV th World Congress of Applied Linguistics, Madison, WI, 173'.

Markham, D. (1997), *Phonetic Imitation, Accent, and the Learner*, Lund: Lund University Press.

Mayer, J. (1995), 'Transcription of German intonation - the Stuttgart System', Technical report, University of Stuttgart.

McAllister, R. (2001), 'Experience as a factor in L2 phonological acquisition', Technical report, Lund University, Dept of Linguistics.

McCarthy, J. (2001), A prosodic theory of nonconcatenative morphology, *in* Charles W. Kreidler, ed., 'Phonology. Critical Concepts', Routledge, London & New York, , pp. 426-477.

McCarthy, J. & Prince, A. (1993), 'Prosodic Morphology: Constraint Interaction and Satisfaction', Technical report, University of Massachusetts and Rutgers University.

McCarthy, P. & Prince, A. (1986), 'Prosodic Morphology'.

McLaughlin, B. (1987), *Theories of Second-Language Learning*, Edward Arnold.

Meara, P. (2005), *LLAMA language aptitude tests*, Swansea. UK: Lognostics.

Mennen, I. (2006), 'Phonetic and phonological influences in non-native intonation: an overview for language teachers'.

Michas, I., & Berry, D. C. (1994), 'Implicit and explicit processes in asecond language learning task', *European Journal of Cognitive Psychology* **6**, 357-381.

Mitchell, R. & Myles, F. (2004), *Second Language Learning Theories*, A Hodder Arnold Publication.

Moulton, W. (1962), *The Sounds of English and German*, University of Chicago Press.

Moyer, A. (2004), *Age, Accent and Experience in Second Language Acquisition. An Integrated Approach to Critical Period Inquiry*, Clevedon: Multilingual Matters Ltd.

Moyer, A. (1999), 'Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction', *Studies in Second Language Acquisition* **21**, 81-108.

Munro, M. (1995), 'Nonsegmental Factors in Foreign Accent: Ratings of Filtered Speech', *Studies in Second Language Acquisition* **17**, 17-34.

Myers, S. (1997), 'Expressing Phonetic Naturalness in Phonology', *Derivation and Constraints in Phonology*, 125-152.

Möbius, B. (1995), Components of a quantitative model of German intonation, *in* 'Proceedings of the 13th International Congress of Phonetic Sciences (Stockholm)', 108-115.

Möbius, B. (1993), *Ein quantitatives Modell der deutschen Intonation*, Tuebingen: Niemeyer.

Möbius, B., Zimmermann, A., Hess, W. (1987), Untersuchungen zu mikroprosodischen Grundfrequenzvariationen im Deutschen, In Hans G. Tillmann and Gerd Willée (eds.), *Analyse und Synthese gesprochener Sprache*, Olms, Hildesheim, 102-110.

Möbius, B. (1985), 'Untersuchungen zur phonetischen Manifestation der phonologischen 'Töne' im Deutschen', Magisterarbeit, Universität Bonn.

Möhler, G. (2001), 'Improvements of the PaIntE model of *F0* parametrization: Technical Report', Technical report, Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

Möhler, G. (1998), 'Theoriebasierte Modellierung der deutschen Intonation für die Sprachsynthese', PhD thesis, Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

Möhler, G. & Conkie, A. (1998). *Parametric modeling of intonation using vector quantization.* Proceedings of 3rd ESCA Workshop on Speech Synthesis, Jenolan Caves, Australia.

Neufeld, G. (1978), 'On the Acquisition of Prosodic and Articulatory Features in Adult Language Learning', *Canadian Modern Language Review* **34**, 163-174.

Neufeld, G. (1978), 'Towards a theory of language learning ability', *Language Learning* **29**, 227-241.

Nosofsky, R. M. (1996), 'Attention, Similarity, and the Identification-Categorization Relationship', *Journal of Experimental Psychology* **115**(1), 39-57.

Nosofsky, R. M. & Zaki, S. (2002), 'Exemplar and Prototype Models Revisited: Response Strategies, Selective Attention, and Stimulus Generalization', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **28**(5), 924-940.

O'Connor, J.D. & Arnold, G. F. (1971), *Intonation of Colloquial English*, London: Longman.

Odlin, T. (1989), *Language Transfer: Cross-Linguistic Influence in Language Learning*, Cambridge University Press.

Oller, J. & Andersen, R. W., ed., (1981), *New dimensions in second language acquisition research*, Rowley, MA: Newbury House, chapter Research on the measurement of affective variables: Some remaining questions, 14-27.

Öhmann, S. (1967), 'Word and sentence intonation: A quantitative model', *STL-QPSR* **8/2-3**.

Pallier, C., Bosch, L. & Sebastian-Gallés, N. (1997), 'A limit on behavioral plasticity in speech perception', *Cognition* **64(3)**, B9-B17.

Palmer, H. (1922), *English Intonation, With Systematic Exercises*, Cambridge: Heffer.

Parry, T.S., Child, J., Parry, T.S. & Stansfield, C., ed., (1990), *Language aptitude reconsidered*, NJ: Prentice Hall, chapter Preliminary investigation of the relationship between VORD, MLAT, and language proficiency.

Peppe, S., M. J. &. W. B. (2000), 'Prosodic Variation in Southern British English', *Language and Speech* **43**, 309-334.

Petersen, C. R. & Al-Haik, A. R. (1976), 'The development of the Defense Language Aptitude Battery (DLAB)', *Educational and Psychological Measurement* **36**, 369-380.

Pierrehumbert, J. Bybee, J., H. P., ed., (2001), *Frequency Effects and the Emergence of Linguistic Structure*, John Benjamins, Amsterdam, chapter Exemplar dynamics: Word frequency, lenition, and contrast, 137–157.

Pierrehumbert, J. (2003), 'Phonetic diversity, statistical learning, and acquisition of phonology', *Language and Speech* **46**(2-3), 115-154.

Pierrehumbert, J. (2000), Tonal elements and their alignment'Prosody: Theory and Experiment. Studies Presented to Gosta Bruce', Kluwer, Dordrecht, 11-26.

Pierrehumbert, J. (2000), 'The phonetic grounding of phonology', *Bulletin de la Communication Parlee* **5**, 7-23.

Pierrehumbert, J. (1980), 'The phonology and phonetics of English intonation', PhD thesis, MIT.

Pimsleur, P. (1966), *The Pimsleur Language Aptitude Battery*, New York: Harcourt Brace Johanovich.
Pinker, S. (1994), *The Language Instinct: How the Mind Creates Language*, New York: Harper Collins.

Piske, T., MacKay, I. & Flege, J. (2001), 'Factors affecting degree of foreign accent in an L2: a review', *Journal of Phonetics* **29**, 191-215.

Pitrelli, J., Beckman, M. & Hirschberg, J. (1994), Evaluation of Prosodic Transcription Labeling Reliability in The ToBI Framework, *in* 'Proceedings of the Third International Conference on Spoken Language Processing (ICSLP 94)', 123-126.

Portele, T. & Heuft, B. (1997), 'Towards a prominence-based speech synthesis system', *Speech Communication* **21**, 61-72.

Poulisse, N. & Bongaerts, T. (1994), 'First Language Use in Second Language Production', *Applied Linguistics* **15-1**, 36-57.

Prince, A. & Smolensky, P. (1993), 'Optimality Theory: Constraint Interaction in Generative Grammar', Technical report, Rutgers University.

Purcell, E.T. & Suter, R. (1980), 'Predictors of pronunciation accuracy: A reexamination', *Language Learning* **30**, 271-287.

Radel, M. (2008), *The Intonation of Declarative and Interrogative Sentences in L2 Spanish by L1 Speakers of German*, Books on Demand GmbH.

Ramirez Verdugo, D. M. (2003), 'Non-native interlanguage intonation systems: A study based on computerized corpus of Spanish learners of English', *ICAME*, 115-132.

Rapp, S. (1998), 'Automatisierte Erstellung von Korpora für die Prosodieforschung', PhD thesis, Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

Rasier, L. & Hiligsmann, P. (2007), 'Prosodic Transfer from L1 to L2. Theoretical and methodological issues'.

Reyelt, M., Grice, M., Benzmüller, R., Mayer, J. & Batliner, A. (1996). Prosodische Etikettierung des Deutschen mit ToBI. Proc. KONVENS 96, 290-299.

Rosch, E. (1999), Reclaiming Concepts, *in* W.J. Nunez, R. & Freeman, ed., 'Reclaiming cognition: The primacy of action, intention and emotion', Thorverton, Eng.: Imprint Academic, .

Rosch, E. (1977), Human categorization, *in* N. Warren, ed., 'Studies in cross-cultural psychology', London: Academic Press, 1-49.

Ross, S., Yoshinaga, N. & Sasaki, M. Robinson, P., ed., (2002), *Individual Differences and Instructed Language Learning*, Amsterdam: Benjamins, chapter Aptitude-exposure interaction effects on Wh-movement violation detection by pre- and post-critical period Japanese bilinguals, 267-299.

Sasaki, M. (1996), *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*, Bern: Peter Lang.

Saussure, F. Séchehaye, C. B. &. A., ed. (1916), *Cours de linguistique générale*, Lausanne/Paris: Payot.

Schachter, P. (1986), 'Review of Foley, William A. and Robert D. Van Valin, Jr., Functional Syntax and Universal Grammar', *Lingua* **69**, 172-186.

Schneider, K. & Möbius, B. (2005), Perceptual Magnet Effect in German boundary tones, *in* 'Interspeech/Eurospeech', 41-44.

Schneiderman, E. I. & Desmarais, C. (1988), 'The talented language learner: Some preliminary findings', *Second Language Research* **4**, 91-109.

Schumann, J. (1978), *The pidginization process: A model for second language acquisition*, New York: Newbury House.

Schumann, J. (1976), 'Affective factors and the problem of age in second language acquisition', *Language Learning* **25**, 209-235.

Schwartz, B. D. (1997), 'On the basis of Basic Variety', *Second Language Research* **13**, 386-402.

Schweitzer, A. (2011), 'Production and Perception of Prosodic Events - Evidence from Corpus-based Experiments', PhD thesis, Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

Schweitzer, A. & Möbius, B. (2009), Experiments on Automatic Prosodic Labeling, *in* 'Proceedings of Interspeech 2009 (Brighton, UK)'.

Schweitzer, A. & Möbius, B. (2004), Exemplar-Based Production of Prosody: Evidence from Segment and Syllable Durations, *in* 'Speech Prosody 2004, International Conference; Nara, Japan, March 23-26, 2004', 459-462.

Schweitzer, K., Riester, A., Walsh, M. & Dogil, G. (2009), Pitch Accents and Information Status in a German Radio News Corpus, *in* 'Proceedings of Interspeech 2009 (Brighton)'.

Schweitzer, K., Walsh, M., Möbius, B., Riester, A., Schweitzer, A. & Schütze, H. (2009), Frequency Matters: Pitch accents and Information Status, *in* 'Proceedings of EACL 2009 (Athens)', 728-736.

Scovel, T. (2000), *Learning New Languages: A Guide to Second Language Acquisition*, Boston: Heinle & Heinle.

Scovel, T. (1988), *A time to speak. A psycholinguistic inquiry into the critical period for human speech*, Rowley, MA: Newbury House.
Scuffil, M. (1982), *Experiments in Comparative Intonation: A Case-Study of English and German*, Berlin, New York: De Gruyter.

Seliger, H. W. & Ritchie, W. C., ed., (1978), *Second language acquisition research: Issues and implications*, San Diego, CA: Academic Press, chapter Implications of a multiple critical periods hypothesis for second language learning, 11-19.

Selinker, H. (1991), *Rediscovering Interlanguage*, Addison Wesley Publishing Company.

Selinker, L. (1972), 'Interlanguage', *IRAL* **10**, 209-231.

Selinker, L., S. M. & Dumas, G. (1975), 'The Interlanguage hypothesis extended too children', *Language Learning* **25**, 139-151.

Selkirk, E. Baayen, H., C. K. & others, ed., (2000), *Prosody: Theory and Experiment*, Dordrecht / Boston / London : Kluwer Academic Publishers, chapter The Interaction of Constraints on Prosodic Phrasing, 231-261.

Selkirk, E. (1995), Sentence Prosody: Intonation, Stress and Phrasing, *in* John Goldsmith, ed., 'The Handbook of Phonological Theory', Oxford: Blackwell, 550-569.

Shafran, I., Ostendorf, M. & Wright, R.(2001), 'Prosody and phonetic variability: Lessons learned from acoustic model clustering', 'Prosody-2001', 127-131.

Sharwood Smith, M. (1994), *Second Language Learning: Theoretical Foundations*, London: Longman.

Shcherba, L. (1974), *Iazykovaia sistema i rechevaia deioiel'nost'*, Leningrad.

Shen, Y. & Peterson, G. (1962), 'Isochronism in English', Technical report, University of Buffalo Studies in Linguistics, Occasional Papers, 9.

Skehan, P. & Robinson, P., ed., (2002), *Individual differences and instructed language learning*, Amsterdam: Benjamins, chapter Theorising and updating aptitude, 69-94.

Skehan, P. (1991), 'Individual Differences in Second Language Learning', *Studies in Second Language Acquisition* **13**, 275-298.

Skehan, P. (1986), 'The role of foreign language aptitude in a model of school learning', *Language Testing* **3**, 188-221.

Snow, C. & Hoefnagel-Hohle, M. (1978), 'The critical period for language acquisition: Evidence from second language learning', *Child Development* **49**, 1114-1128.

Snow, C. & Hoefnagel-Hohle, M. (1977), 'Age differences in the pronunciation of foreign sounds', *Language and Speech* **20**, 357-365.

Sparks, R. L., Artzer, M., Ganschow, L., Siebenhar, D., Plageman, M. & Patton, J. (1998), 'Differences in native-language skills, foreign-language aptitude, and foreign-language grades among high-, average-, and low-proficiency foreign-language learners: two studies', *Language Testing* **15**, 181-216.

Sparks, R., Javorsky, J., Patton & Ganschow, L. (1998), 'Factors in the Prediction of Achievement and Proficiency in a Foreign Language', *Applied Language Learning* **9**, 71-105.

Sweet, H. (1878), *Handbook of phonetics*, Oxford: Clarendon Press.

Tarone, E. (1983), 'On the Variability of Interlanguage Systems', *Applied Linguistics* **4-2**, 142-164.

Tarone, E. (1988), *Variation in Interlanguage*, Edward Arnold.

Tarone, E. (1979), 'Interlanguage as Chameleon', *Language Learning* **29-1**, 181-191.

Tarone, E. (1978), The Phonology of Interlanguage'Understanding Second and Foreign Language Learning', Rowley, MA.

Taylor, J. R. (1990), Schemas, prototypes, and models: in search of the unity of the sign, *in* S.L. Tsohatzidis, ed., 'Meanings and Prototypes. Studies in Linguistic Categorization', Routledge, London & New York, 521-534.

Taylor, P. (2000), 'Analysis and Synthesis of Intonation using the Tilt Model', Technical report, Centre for Speech Technology Research, University of Edinburgh.

Towell, R. & Hawkins, R. (1993), 'Systematic and Nonsystemativ Variability in Advanced Language Learning', *Studies in Second Language Acquisition* **15**, 439-460.

Trager, G.L. & Smith, H. (1951), 'An outline of English structure', *Studies in Linguistics occasional papers* **3**.

Trofimovich, P. (2005). Spoken-word processing in native and second languages: An investigation of auditory word priming. *Applied Psycholinguistics, 26*, 479-504.

Trofimovich, P. & Baker, W. (2006), 'Learning Second Language SuprasegmentalsL: Effect of L2 Experience on Prosody and Fluenc Characteristics of L2 Speech', *Studies in Second Language Acquisition* **28**, 1-30.

Trubetzkoy, N.S. (1958), *Grundzüge der Phonologie,* Göttingen: Vandenhoeck & Ruprecht.

Uhmann, S. (1991), *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*, Tuebingen: Niemeyer.

Uhmann, S. (1988), Akzenttöne, Grenztöne und Fokussilben. Zum Aufbau eines phonologischen Intonationssystems für das Deutsche'Intonationsforschungen', Tuebingen: Niemeyer.

Van Els, T. & De Bot, K. (1987), 'The Role of Intonation in Foreign Accent', *The Modern Language Journal* **71-2**, 147-155.

Van Santen, J. & Hirschberg, J. (1994), Segmental effects on timing and height of pitch and contours, *in* 'Proceedings of ICLSP '94 (Yokohama, Japan)'.

Van Santen, J. & Möbius, B. (2000), A Quantitative Model of F0 Generation and Alignment'Intonation: Analysis, Modelling and Technology', Dordrecht / Boston / London: Kluwer Academic Publishers.

Vandeloise, C. (1990), Representation, prototypes, and centrality, *in* S.L. Tsohatzidis, ed., 'Meanings and Prototypes. Studies in Linguistic Categorization', Routledge, London & New York, 403-437.

Vella, A. (1995), 'Prosodic Structure and Intonation in Maltese and its Influence on Maltese English', PhD thesis, University of Edinburgh.

Walsh, T. & Diller, K. Diller, K., ed., (1981), *Individual differences and universals in language learning aptitude*, Rowley, MA: Newbury House, chapter Neurolinguistic considerations on the optimum age for second language learning, 3-21.

Weiner, B. (1992), *Human Motivation: Metaphors, Theories, and Research*, Sage Publications.

Wesche, M. B. Diller, K., ed., (1981), *Individual differences and universals in language learning aptitude*, Rowley, MA: Newbury House, chapter Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems, 119-139.

Willems, N., C. R. 't. H. J. (1988), 'A synthesis scheme for British English intonation', *Journal of the Acoustical Society of America* **84**, 1250-1261.

Wunderlich, D. (1988), Der Ton macht die Melodie - Zur Phonologie der Intonation des Deutschen'Intonationsforschungen', Tuebingen: Niemeyer.

CURRICULUM VITAE

ANUFRYK Olga Volha

Olga.Anufryk@ims.uni-stuttgart.de

Work experience

10.2009 - present time:	TTS Voice Developer at Nuance Communications, Ulm, Germany
10.2008 - 02.2009:	Lecturer at the Department of English, Universität Stuttgart - Seminar "Foreign Accent in English: Didactics and Instrumental Analysis"
09.2006 - 03.2007:	German / English lecturer at Belarusian State Technical University
09.2004 - 06.2005:	Lecturer in Practical English Phonetics, Reading and Listening Comprehension at Minsk State Linguistic University
01.2005 - 09.2006:	Customer Service Manager at Printcorp Printing and Publishing, Minsk, Belarus
09.2004 - 01.2005:	Interpreter at Printcorp Printing and Publishing, Minsk, Belarus

Education and training

04.2007 - present time:	PhD studies at the Graduate School 609 "Language Representation and their Interpretations", Universität Stuttgart, Institute for Natural Language Processing
09.2004 - 06.2005:	Master of Philological Sciences (speciality: linguist, university lecturer / teacher) from Minsk State Linguistic University

09.1999 - 06.2004:	Bachelor of Arts in Linguistics (speciality: linguist, teacher of English, German and Rhetoric) from Minsk State Linguistic University
10.2000 - 06.2003:	French Language Proficiency Certificate (speciality: teacher and interpreter / translator of French) from Minsk State Linguistic University

Awards

01.2007:	PhD Grant from the Graduate School 609 "Linguistic Representations and Their Interpretation"
04.2005:	First-degree award at the Belarusian National Contest of Students' Scientific Papers in Humanities

Skills and competences

Languages:

Russian, Belarusian:	Mother tongues
English:	Native-like proficiency
German:	Advanced-level proficiency
French:	Intermediate proficiency
Polish:	Intermediate proficiency
Italian:	Basic-level skills

Programming:

Perl, AWK, HTML

Operating systems and software:

Linux, Unix, Windows

SPSS, R

International language exams:

TOEFL:	693 points on the PBT
GRE:	83 % Verbal, 60 % Quantitative, 75 % Writing
TestDAF:	"5" in all sections

Curriculum Vitae

APPENDIX

The current Appendix contains all illustration and statistical materials not included in the text of the dissertation. They do, however, serve as detailed support to the observations laid down in Sections 7.1.1-4.

A few comments ought to be made about the layout and content of the materials presented below.

Tables of statistical significance

Thus, it should be noted that in all the tables of statistical significance (Tables 1-30) subject group identities are coded by means of abbreviations:

- HH – highest aptitude
- H – high aptitude
- AA – above-average
- A – average
- BA – below-average
- LA – low-aptitude
- N - native

The statistical significance values in those tables are presented in the form of matrices, where the corresponding *p*-value is to be found at each corresponding crossing cell of two group identities, e.g. H (high-aptitude) and AA (above-average), found in the horizontal and vertical rows listing all groups under investigation split by the investigated factors (speech type – read, spontaneous or whole corpus in general; gender – male or female). Statistically significant contrasts are highlighted in light grey.

L*H accents

Parameter *d*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	1					.574					.000				
	AA	.000	.000				.001	.021				.000	.088			
	A	.621	.113	.000			.000	.000	.313			.000	.001	.000		
	BA	.000	.000	.000	.000		.000	.000	1	.001		.000	.000	.000	.000	
	LA	.000	.000	.019	.000	.000	.000	.000	.000	.000	.000	.001	.000	.000	.751	.000
read speech	HH															
	H	.005					.670					.000				
	AA	.000	.000				.000	.000				.000	1			
	A	.844	.000	.000			.000	.000	.678			.000	.000	.000		
	BA	.003	1	.000	.000		.000	.000	1	.156		.000	.000	.000	.000	
	LA	.000	.000	.068	.000	.000	.000	.000	.000	.000	.000	.000	.013	.009	.473	.000
spontaneous speech	HH															
	H	.000					.748					.108				
	AA	.000	.648				.996	.999				.000	.000			
	A	.969	.000	.000			.007	.360	.676			.456	.825	.000		
	BA	.000	.000	.000	.000		.996	.904	1	.003		.998	.992	.473	1	
	LA	.000	1	.625	.000	.000	.000	.000	.006	.000	.000	.876	.000	.000	.000	.934

*Table 1: Inter-group differences in parameter d of L*H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.830						.996						.000					
	AA	.000	.000					.000	.000					.188	.000				
	A	.001	.000	.000				.000	.000	.957				.999	.000	.000			
	BA	.006	.175	.000	.000			.000	.000	.763	.919			.000	.000	.000	.000		
	LA	.000	.000	.000	.979	.000		.000	.000	.750	.000	.000		1	.000	.001	.990	.000	
	N	.000	.000	.000	.543	.000	.959	.000	.000	.016	.000	.003	.000	.049	.000	1	.000	.000	.000
read speech	HH																		
	H	.385						.911						.000					
	AA	.000	.000					.000	.000					.000	.000				
	A	.000	.000	.000				.000	.000	.986				.984	.000	.000			
	BA	.863	.000	.000	.000			.000	.000	.984	1			.000	.000	.024	.000		
	LA	.000	.000	.000	1	.000		.000	.000	.108	.000	.001		.392	.000	.000	.127	.000	
	N	.000	.000	.000	.011	.000	.011	.000	.000	.007	.000	.000	.000	.350	.000	.003	.173	.000	1
spontaneous speech	HH																		
	H	.908						1						.000					
	AA	.021	.000					.058	.021					.188	.000				
	A	.105	.482	.000				.864	.574	.167				.999	.000	.000			
	BA	.000	.000	.000	.000			.977	.957	.000	.001			.000	.000	.000	.000		
	LA	.399	.950	.000	.932	.000		.017	.000	.963	.000	.000		1	.000	.001	.990	.000	
	N	1	.464	.000	.000	.000	.000	.462	.110	.420	.748	.000	.027	.049	.000	1	.000	.000	1

*Table 2: Inter-group differences in parameter d of L*H accents in the English part of the corpus*

Parameter *a1*

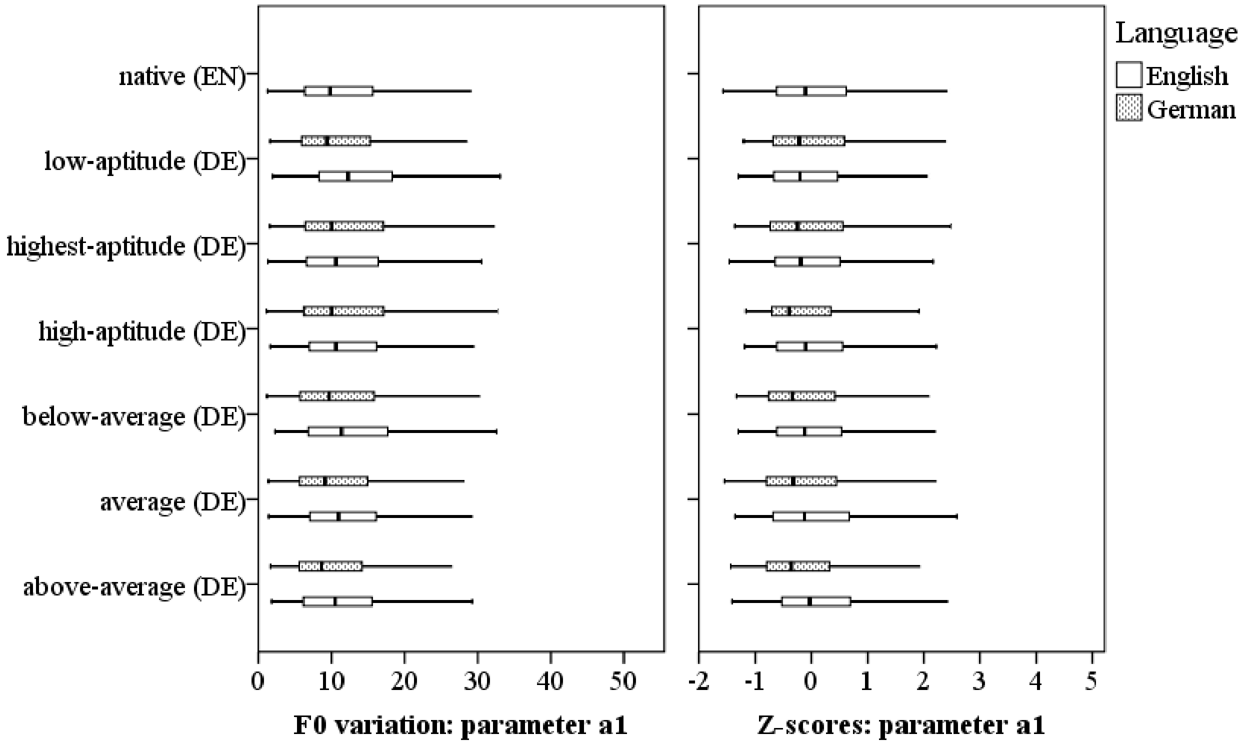


*Figure 1: a1 parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German*

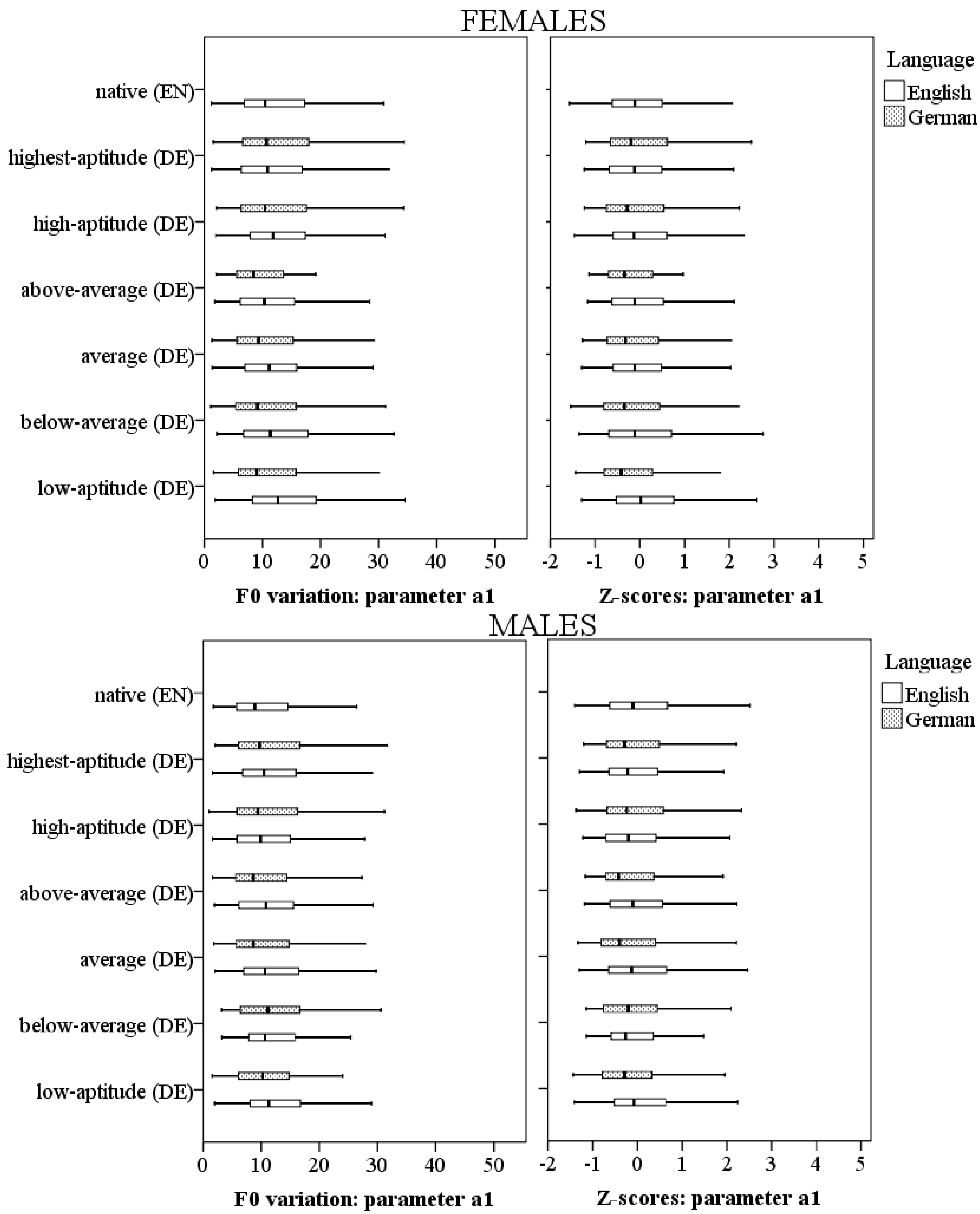


*Figure 2: a1 parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German genderwise*

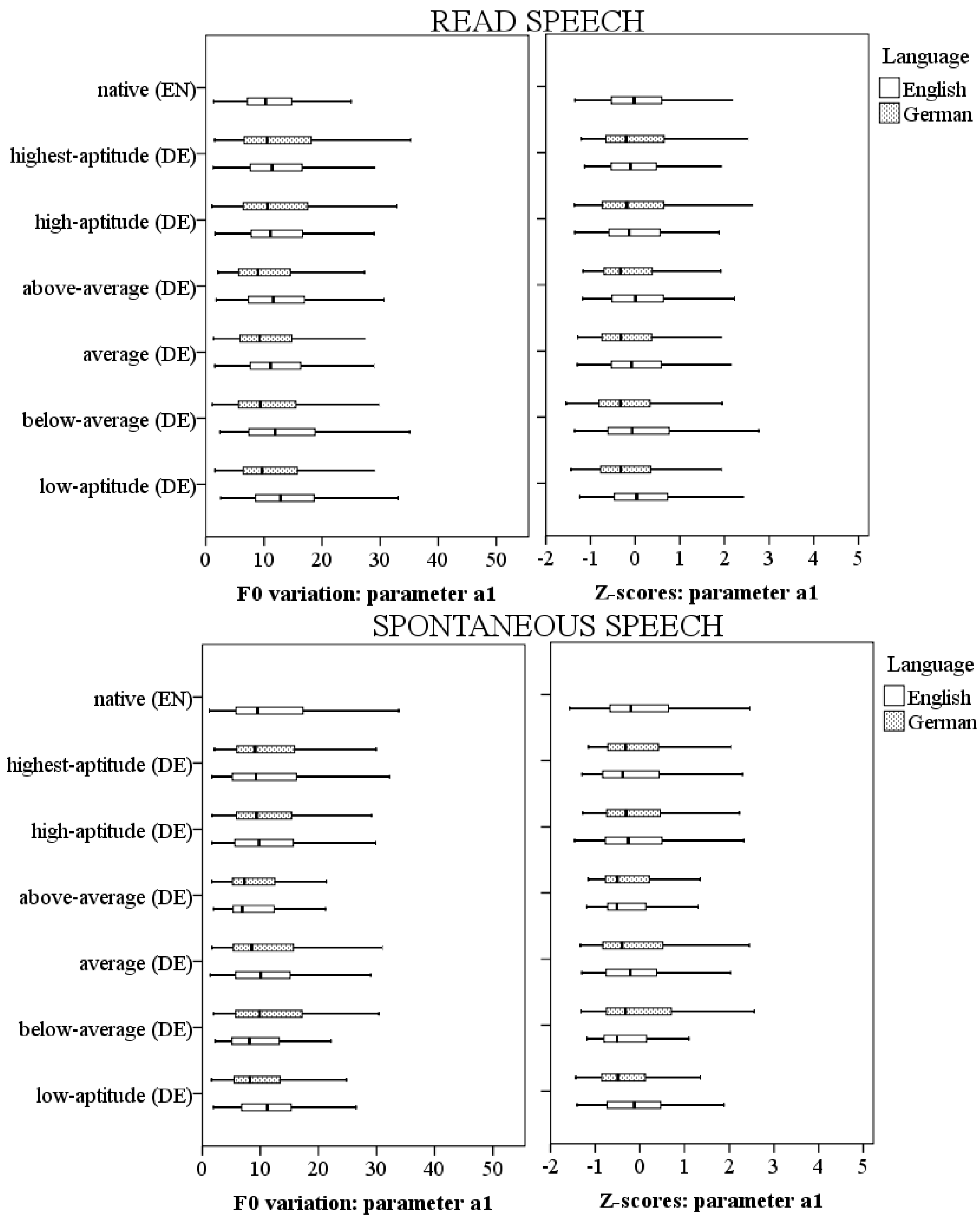


*Figure 3: a1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

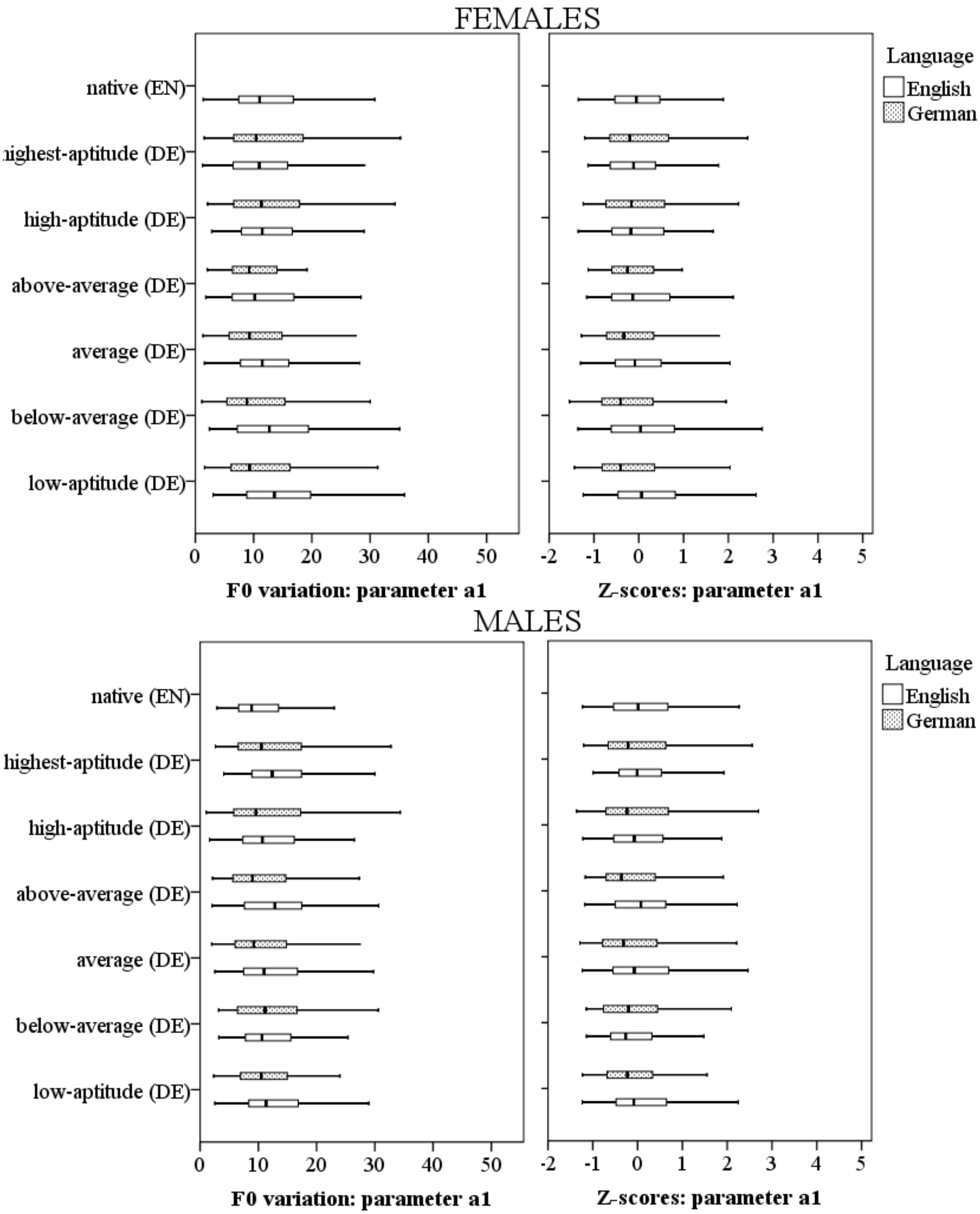


*Figure 4: a1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German read speech by gender – male versus female realizations*

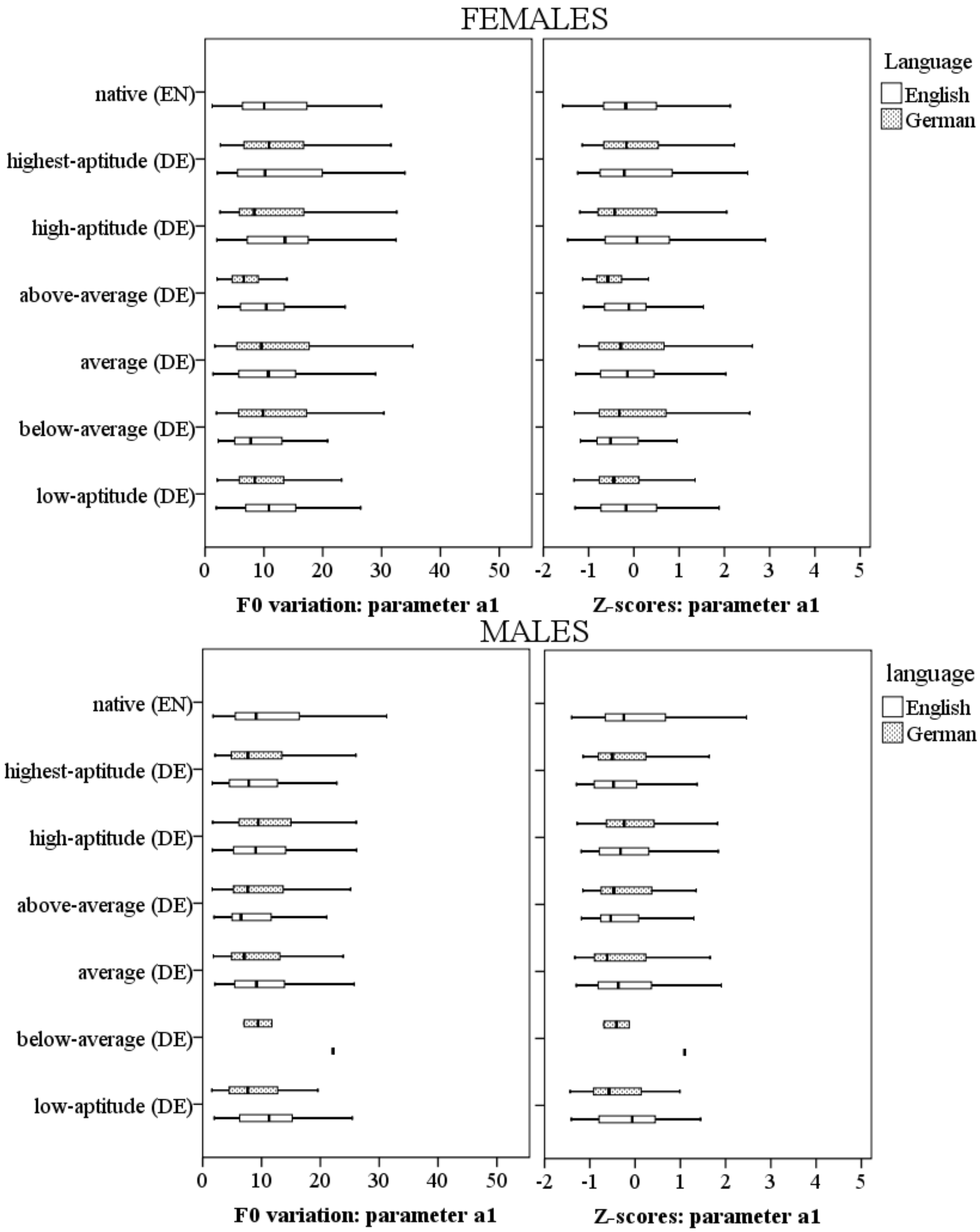


*Figure 5: a1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German spontaneous speech by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.999					1					1				
	AA	.035	.016				.044	.031				.844	.778			
	A	.095	.022	.870			.269	.110	.494			.603	.346	.998		
	BA	.659	.695	.402	.795		.396	.272	.495	1		.999	.988	.593	.319	
	LA	.051	.008	.956	.998	.599	.079	.012	.778	.891	.915	.853	.744	1	.969	.583
read speech	HH															
	H	.970					.996					.976				
	AA	.107	.234				.262	.323				.754	.981			
	A	.012	.010	1			.061	.013	1			.511	.869	1		
	BA	.116	.234	.999	.975		.117	.059	.999	1		1	.994	.845	.613	
	LA	.124	.223	.982	.676	.999	.230	.145	.974	.890	.983	.835	.998	.998	.917	.914
spontaneous speech	HH															
	H	.992					1					.843				
	AA	.644	.098				.105	.069				1	.626			
	A	1	.989	.153			1	.999	.017			.994	.038	.917		
	BA	.817	.951	.021	.653		.993	.983	.015	.997		.999	.935	.997	1	
	LA	.532	.007	1	.003	.001	.527	.278	.485	.001	.011	.992	.027	.892	1	1

*Table 3: Inter-group differences in the parameter a1 of L*H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.992						.976						.593					
	AA	.999	1					.984	.639					1	.375				
	A	1	.868	.984				.999	.994	.778				1	.334	1			
	BA	.997	.618	.868	.986			.997	.999	.737	1			1	.892	1	1		
	LA	.039	.000	.000	.000	.004		.051	.337	.006	.000	.000		.984	.001	.802	.172	.966	
	N	.705	.967	.893	.006	.007	.000	1	.851	.962	.931	.902	.000	.182	.996	.017	.001	.620	.000
read speech	HH																		
	H	.950						1						.346					
	AA	1	.746					1	.989					1	.251				
	A	.997	.988	.909				.989	1	.897				.454	.992	.213			
	BA	.999	.489	1	.562			.694	.834	.472	.677			.384	1	.348	.986		
	LA	.365	.001	.350	.000	.112		.009	.007	.006	.000	.012		.910	.666	.913	.703	.731	
	N	.044	.410	.001	.000	.000	.000	1	1	.993	.966	.261	.000	.000	.032	.000	.000	.261	.000
spontaneous speech	HH																		
	H	.998						.962						.998					
	AA	.517	.055					.975	.560					.980	.531				
	A	.965	.999	.004				1	.621	.963				.924	.982	.106			
	BA	.999	.822	.697	.364			.647	.007	1	.060			.116	.152	.047	.215		
	LA	.591	.674	.000	.730	.026		1	.780	.931	1	.026		.183	.079	.000	.346	.438	
	N	.955	.998	.002	1	.289	.675	1	.550	.970	1	.061	.998	.701	.693	.009	.991	.265	.608

*Table 4: Inter-group differences in the parameter a1 of L*H accents in the English part of the corpus*

Parameter *b*

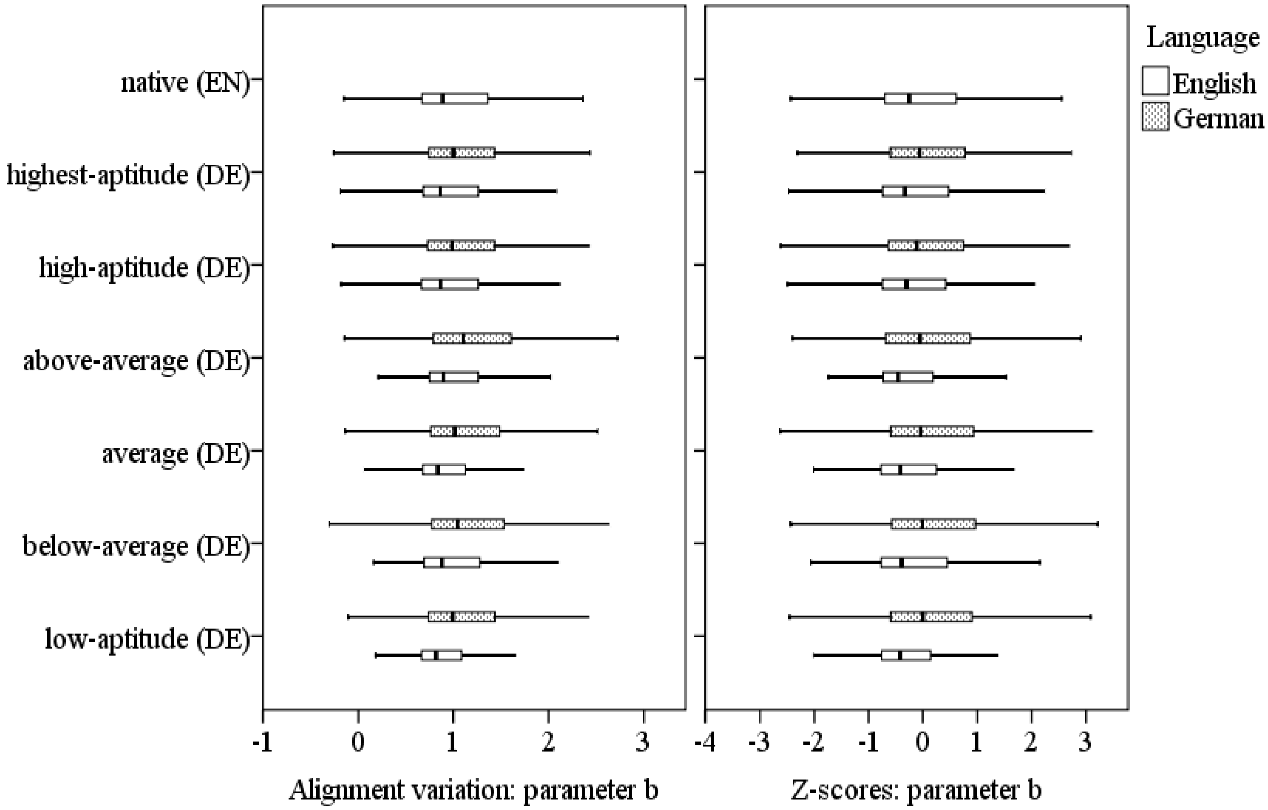


*Figure 6: b parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German*

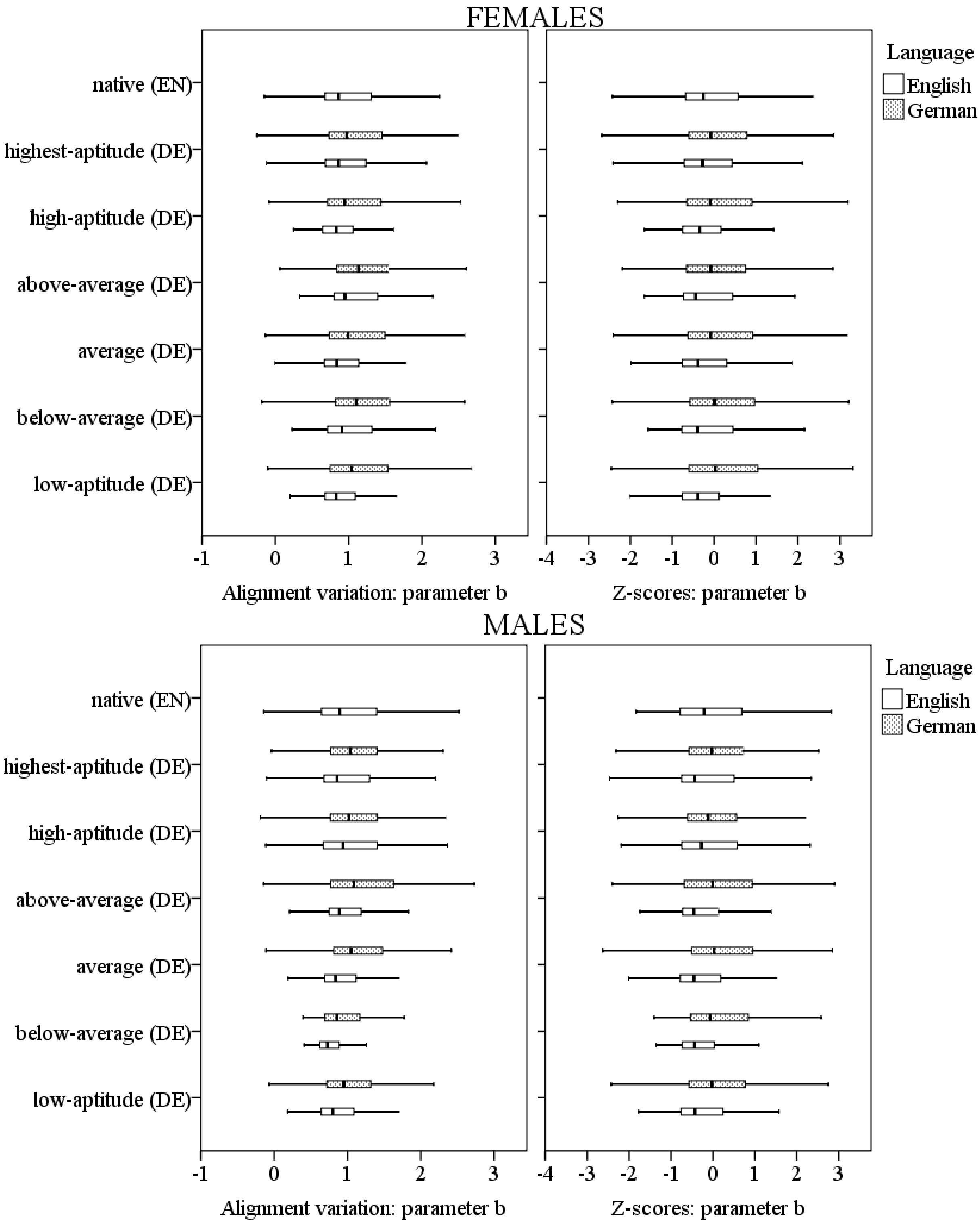


*Figure 7: b parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German genderwise*

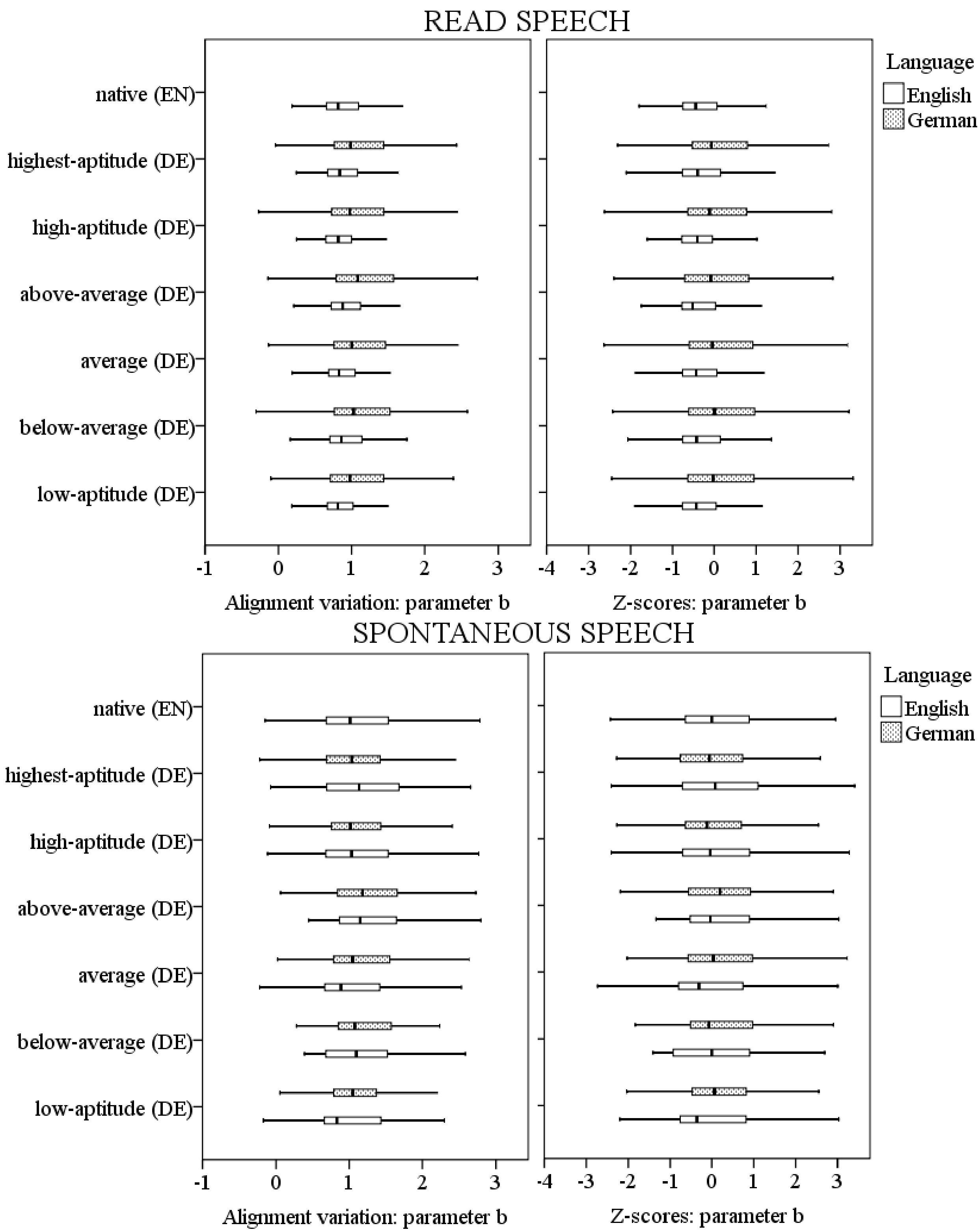


*Figure 8: b parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

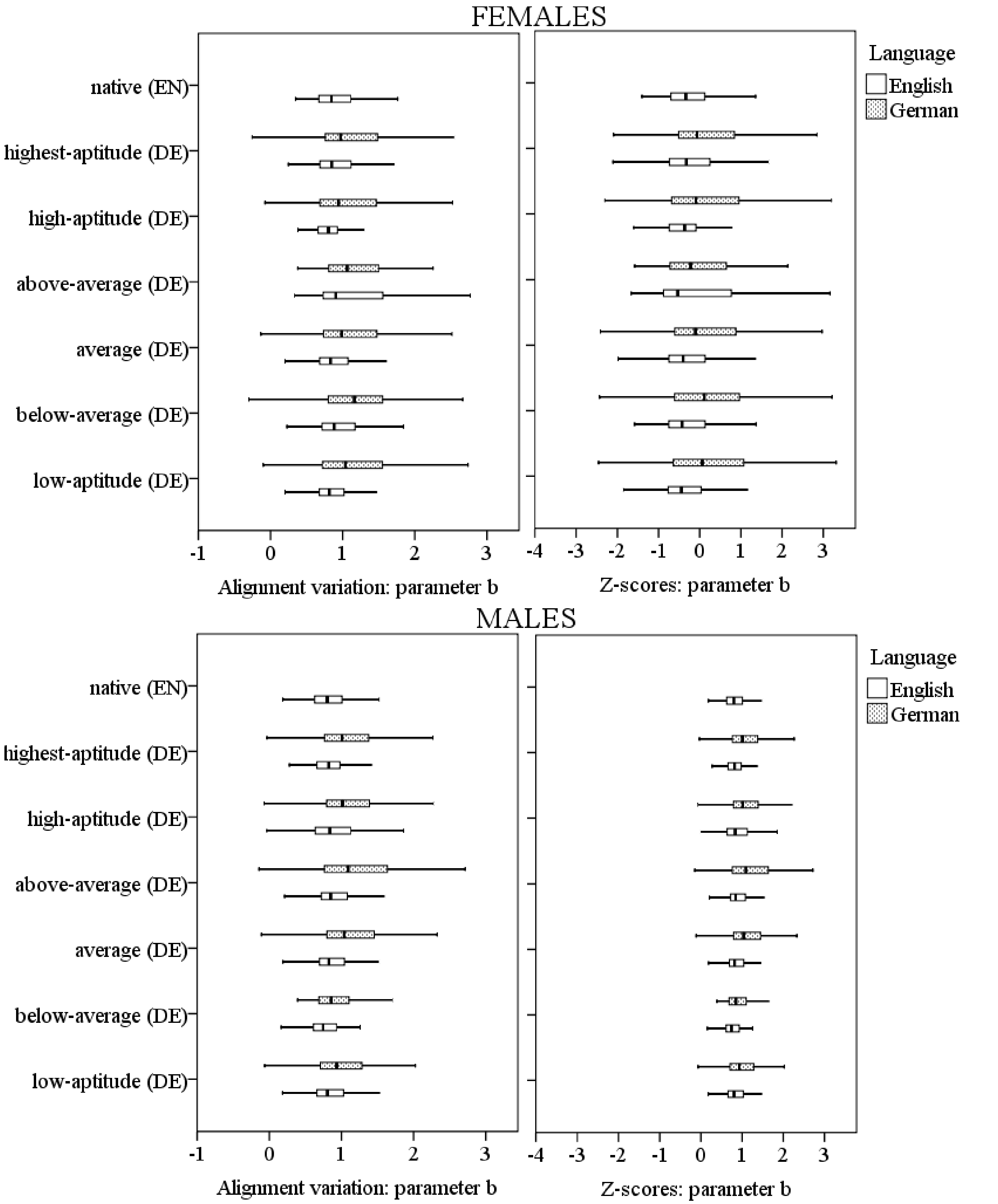


*Figure 9: b parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German read speech by gender – male versus female realizations*

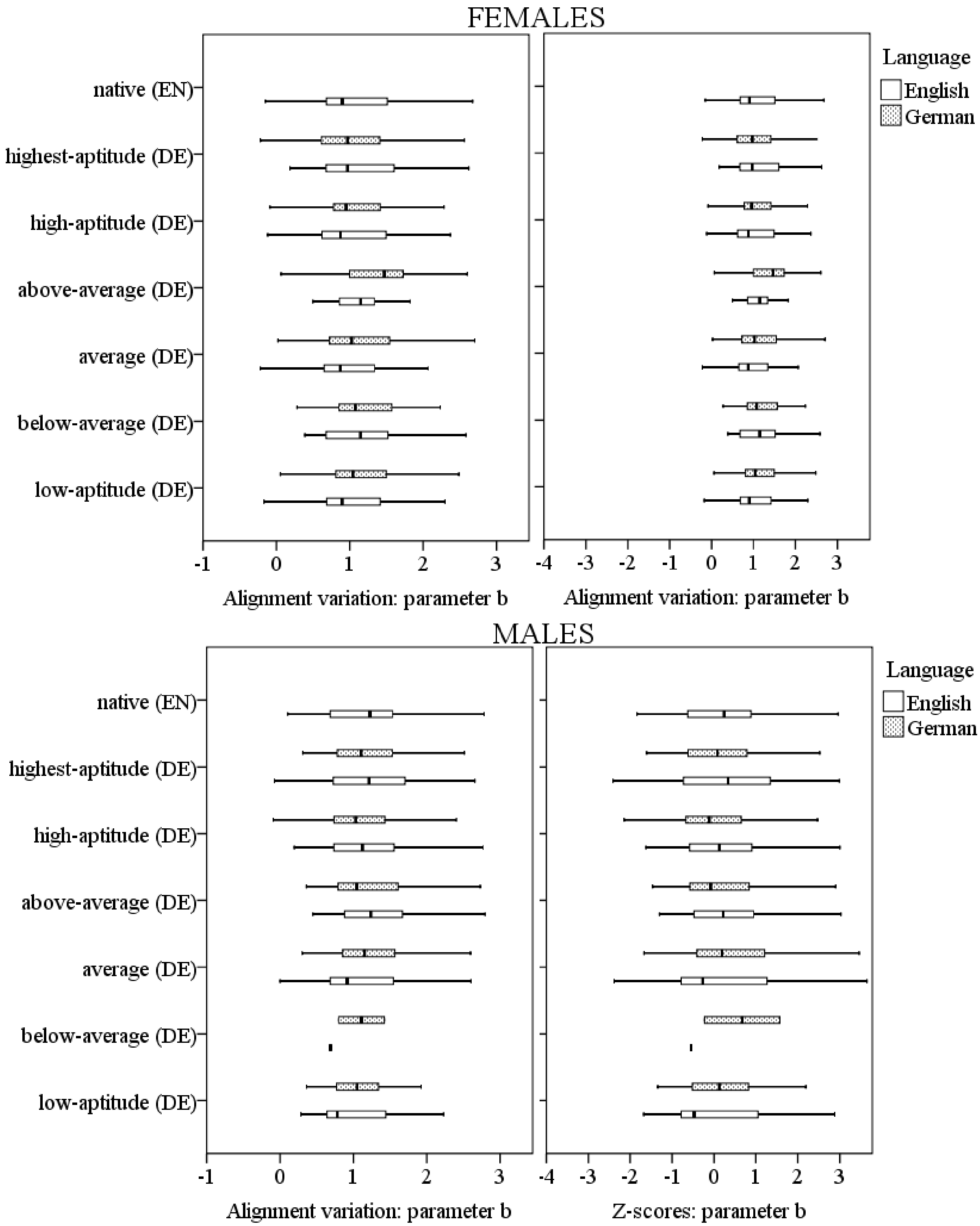


*Figure 10: b parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German spontaneous speech by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	1					1					1				
	AA	.002	.000				.048	.006				.300	.096			
	A	.720	.254	.005			.902	.383	.095			.895	.768	.525		
	BA	.237	.030	.330	.705		.024	.000	.963	.003		.036	.003	.000	.000	
	LA	1	.996	.000	.198	.020	.588	.069	.275	.866	.099	.467	.048	.000	.000	.329
read speech	HH															
	H	.997					.882					.975				
	AA	.113	.002				.914	.248				.126	.302			
	A	.999	.760	.018			1	.731	.648			.878	1	.189		
	BA	.800	.162	.505	.648		.634	.001	1	.006		.172	.003	.000	.000	
	LA	.997	1	.000	.463	.038	.997	.115	.954	.608	.324	.578	.008	.000	.000	.726
spontaneous speech	HH															
	H	.996					.899					.936				
	AA	.056	.053				.001	.011				1	.685			
	A	.278	.260	.711			.181	.848	.038			.994	.114	.978		
	BA	.124	.135	.999	.913		.016	.210	.358	.561		.999	1	.999	.991	
	LA	.771	.898	.207	.750	.423	.272	.929	.028	1	.443	.988	.994	.897	.261	1

*Table 5: Inter-group differences in the parameter b of L*H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	1						.421						.862					
	AA	.498	.194					.109	.000					.997	.977				
	A	.441	.261	.000				.880	.812	.000				.740	.000	.018			
	BA	1	.993	.450	.002			.883	.000	.341	.000			.001	.000	.000	.005		
	LA	.050	.007	.000	.469	.000		.590	.984	.000	.941	.000		.172	.000	.000	.552	.084	
	N	.743	.292	.979	.000	.657	.000	.974	.001	.166	.000	.996	.000	.878	1	.985	.000	.000	.000
read speech	HH																		
	H	.939						.334						.978					
	AA	.392	.005					.010	.000					.894	1				
	A	1	.762	.026				.999	.170	.000				.998	.997	.905			
	BA	.275	.000	1	.000			.570	.000	.130	.000			.453	.032	.004	.022		
	LA	.973	.999	.000	.527	.000		.881	.677	.000	.684	.000		1	.893	.452	.941	.113	
	N	1	.589	.157	.998	.028	.367	1	.074	.000	.990	.033	.364	.999	.998	.938	1	.040	.981
spontaneous speech	HH																		
	H	.995						.950						1					
	AA	.662	.092					1	.874					.931	.477				
	A	.130	.105	.000				.692	.999	.627				.719	.582	.006			
	BA	.998	1	.154	.165			1	.576	1	.012			.661	.708	.410	.878		
	LA	.137	.120	.000	1	.178		.942	1	.865	.916	.223		.164	.030	.000	.750	.959	
	N	.960	1	.009	.005	1	.010	.999	.970	.979	.098	.778	.856	.992	.999	.113	.630	.754	.012

*Table 6: Inter-group differences in the parameter b of L*H accents in the English part of the corpus*

Parameter *c1*

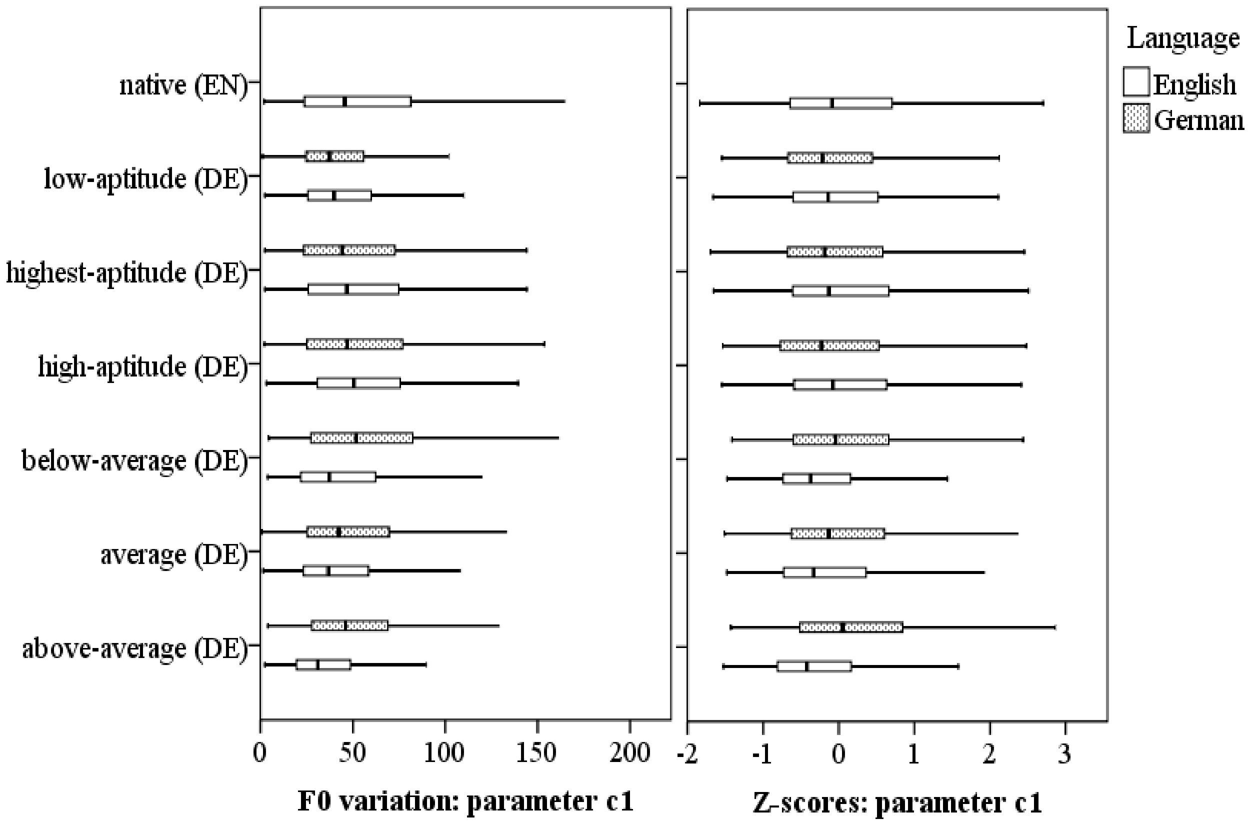


*Figure 11: c1 parameter variation in L*H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German*

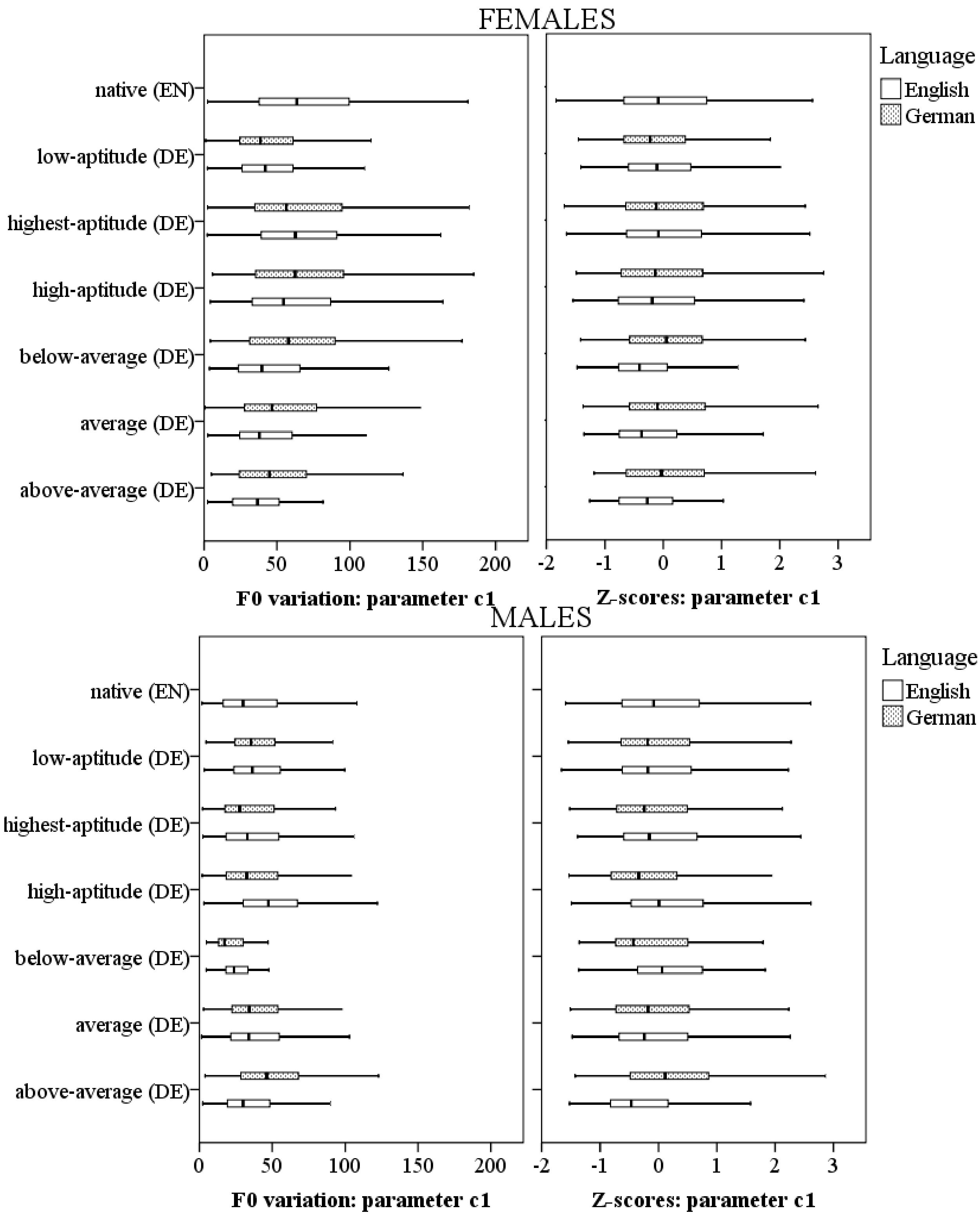


*Figure 12: c1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German genderwise*

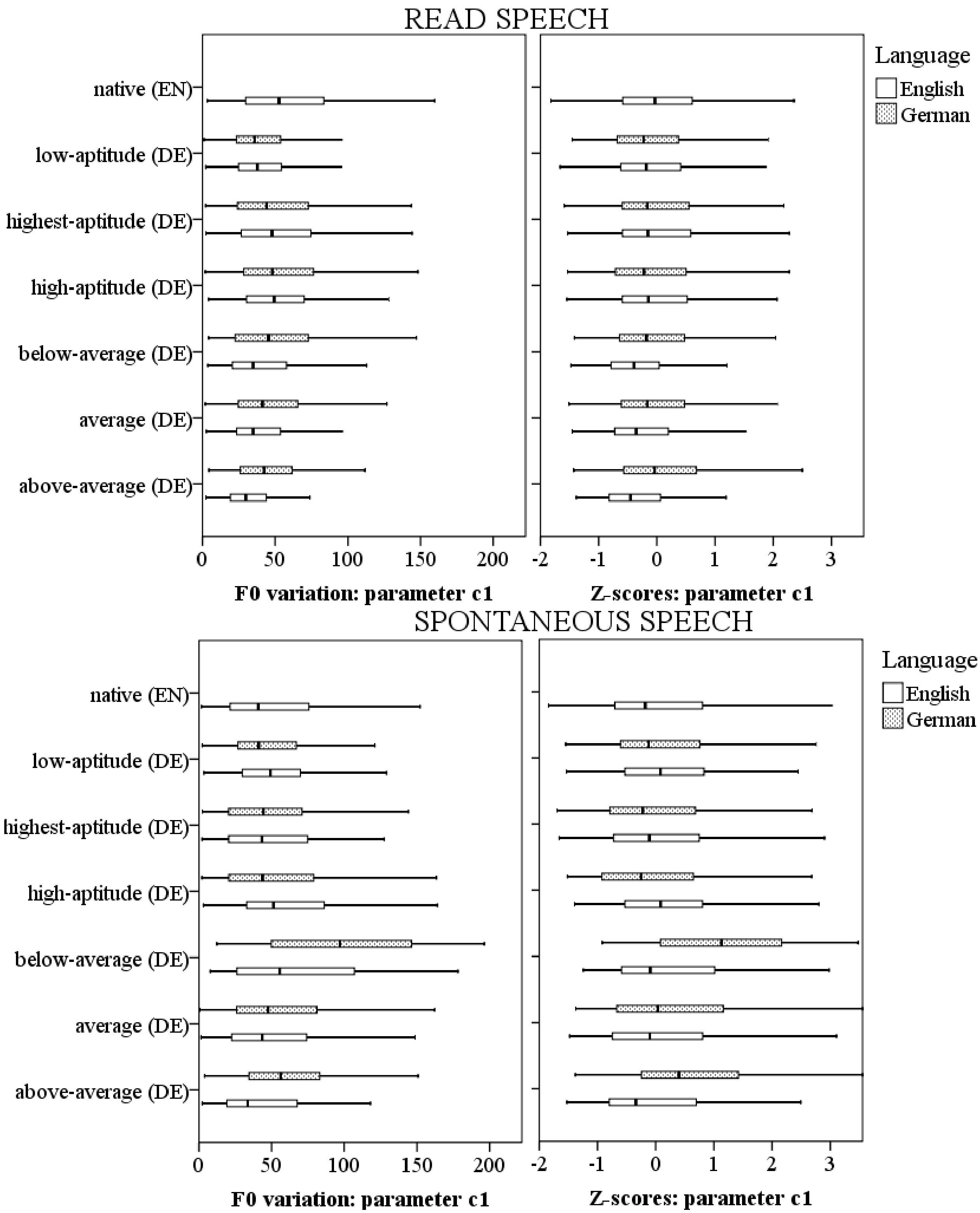


*Figure 13: c1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

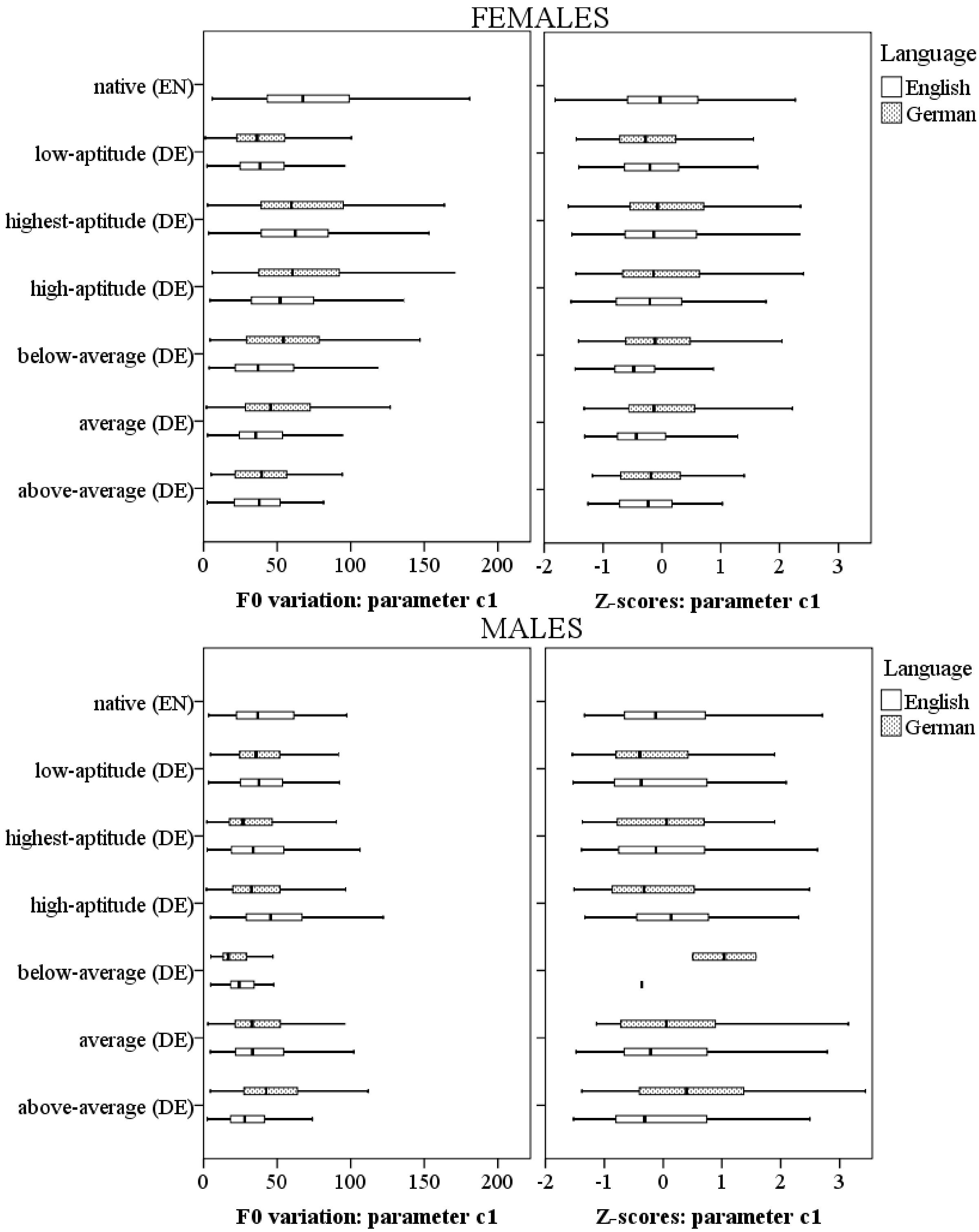


*Figure 14: c1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German read speech by gender – male versus female realizations*

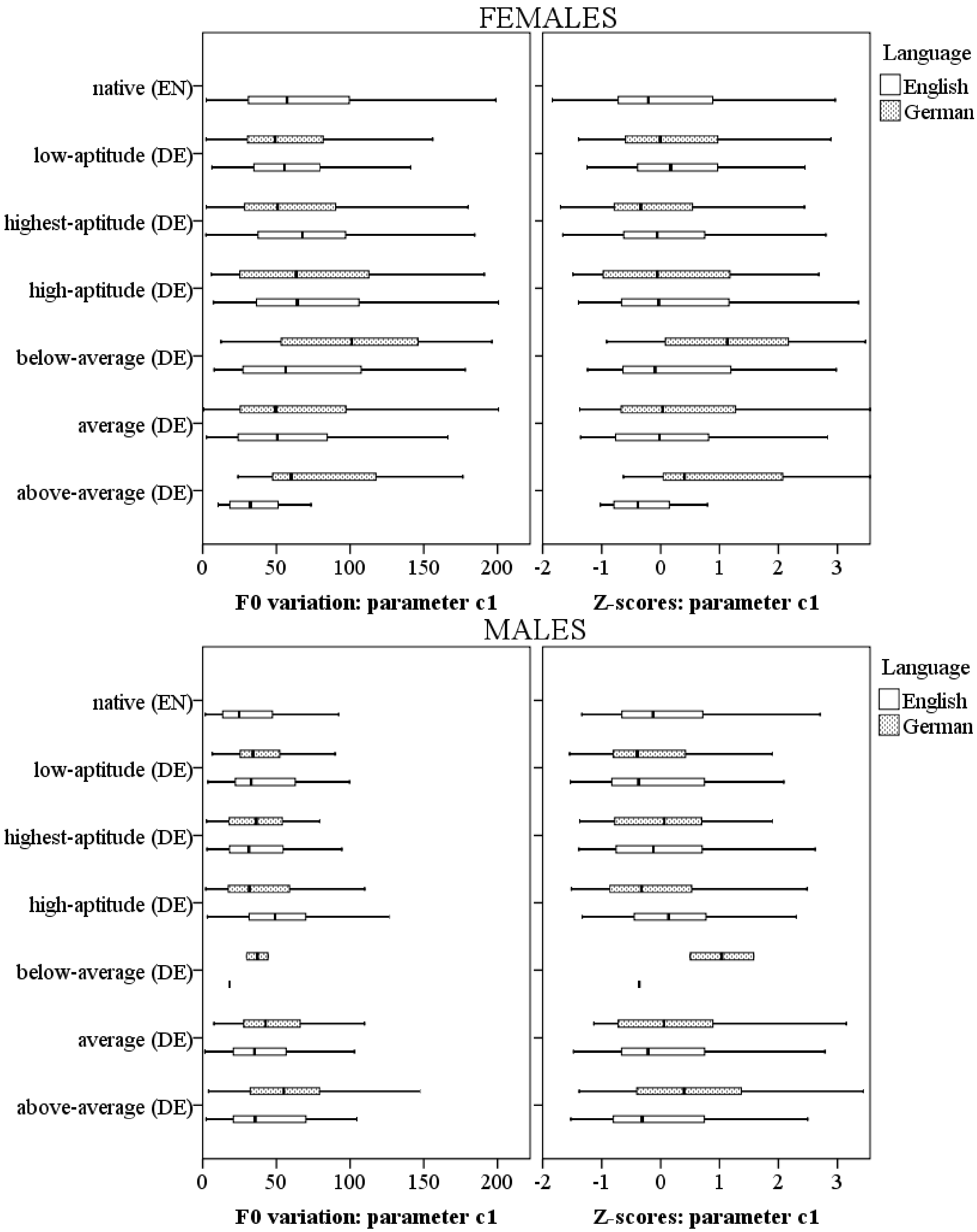


*Figure 15: c1 parameter variation in L*H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German spontaneous speech by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.737					.952					.354				
	AA	.999	.300				.013	.000				.000	.000			
	A	.986	.014	1			.010	.000	.013			.007	.537	.000		
	BA	.006	.093	.000	.000		.999	.979	.000	.000		.000	.000	.000	.000	
	LA	.000	.000	.000	.000	.000	.000	.000	.490	.000	.000	.044	.936	.000	.926	.000
read speech	HH															
	H	.623					1					.793				
	AA	.322	.000				.000	.000				.000	.000			
	A	.461	.000	.971			.000	.000	.213			.098	.728	.000		
	BA	1	.134	.173	.143		.089	.001	.004	.046		.000	.000	.000	.000	
	LA	.000	.000	.013	.000	.000	.000	.000	.828	.000	.000	.081	.670	.000	1	.000
spontaneous speech	HH															
	H	1					.743					.839				
	AA	.387	.332				.754	.997				.000	.000			
	A	.704	.625	.913			1	.402	.672			.131	.349	.002		
	BA	.000	.000	.000	.000		.000	.000	.155	.000		1	.998	.356	.935	
	LA	1	.978	.045	.027	.000	.954	.019	.267	.404	.000	.843	1	.000	.218	.998

*Table 7: Inter-group differences in the parameter c1 of L*H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.742						.979						.000					
	AA	.000	.000					.000	.000					.708	.000				
	A	.000	.000	.000				.000	.000	.098				.798	.000	.000			
	BA	.063	.000	.000	.293			.000	.000	.006	.415			.001	.000	.021	.000		
	LA	.000	.000	.000	1	.280		.000	.000	.107	1	.325		.945	.000	.000	.994	.000	
	N	1	.548	.000	.000	.000	.000	.000	.333	.000	.000	.000	.000	.995	.000	.733	.000	.000	.015
read speech	HH																		
	H	1						.535						.012					
	AA	.000	.000					.000	.000					.099	.000				
	A	.000	.000	.000				.000	.000	.888				.976	.002	.000			
	BA	.000	.000	.000	.956			.000	.000	.213	.138			.002	.000	.453	.000		
	LA	.000	.000	.000	1	.943		.000	.000	.819	1	.229		.998	.000	.000	.994	.000	
	N	.263	.262	.000	.000	.000	.000	.401	.000	.000	.000	.000	.000	.797	.071	.000	.941	.000	.682
spontaneous speech	HH																		
	H	.708						1						.001					
	AA	.484	.000					.041	.005					.784	.053				
	A	1	.161	.094				.628	.074	.248				.912	.000	.997			
	BA	.075	.689	.000	.000			.999	.948	.024	.331			.885	.189	.615	.692		
	LA	.997	.669	.015	.948	.010		.879	.309	.113	.965	.837		.953	.000	.992	1	.716	
	N	.998	.004	.378	.830	.000	.169	.997	.832	.020	.102	1	.731	.866	.000	.001	.000	.970	.000

Table 8: Inter-group differences in the parameter c1 in the English part of the corpus

H* accents

Parameter *b*

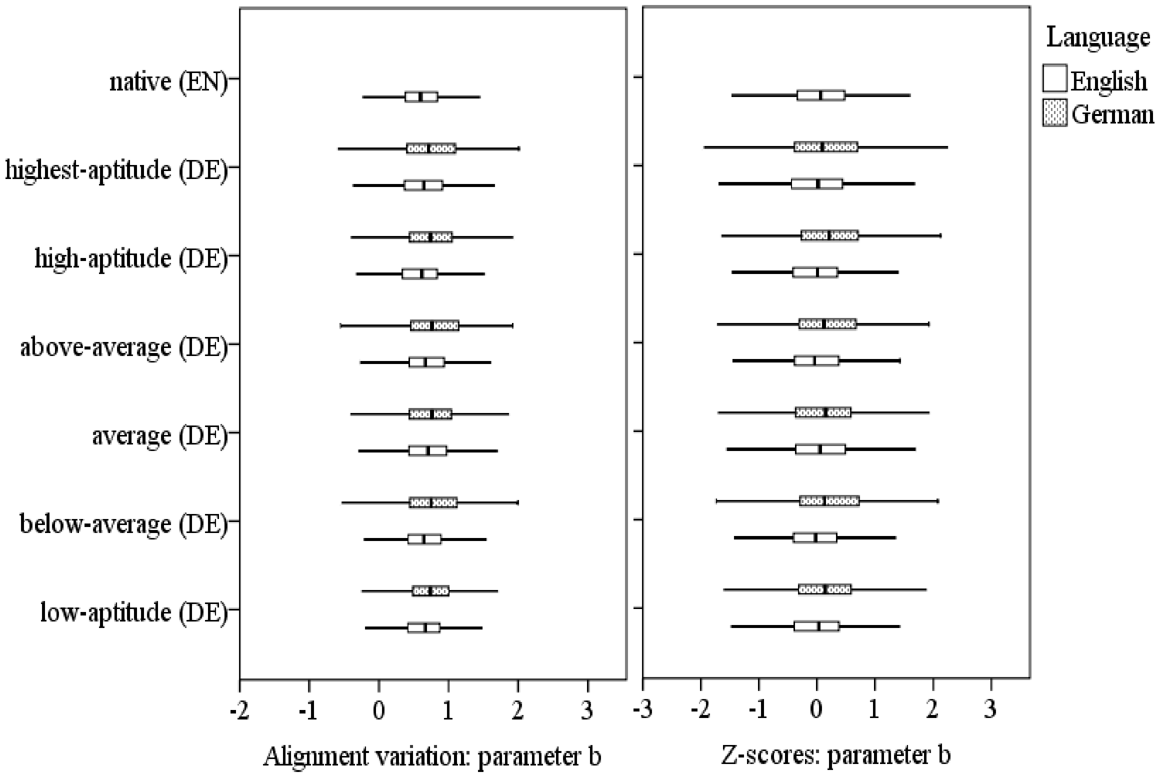


Figure 16: b parameter variation in H accents with regard to the F0 level in Hz (left chart) and the within-speaker / group variability reflected by the z-scores (right chart) in all the aptitude groups in English and German*

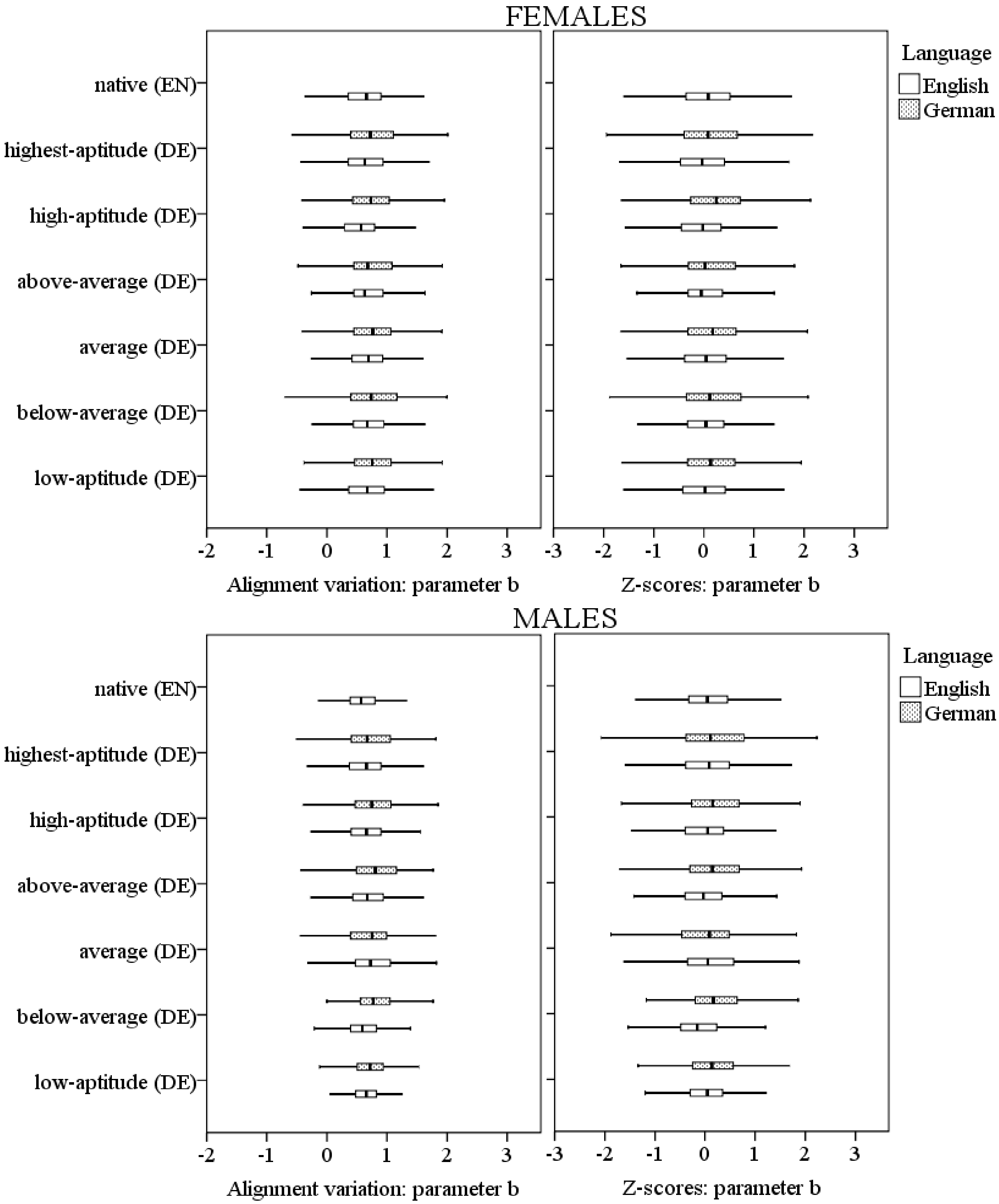


Figure 17: b parameter variation in H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German genderwise*

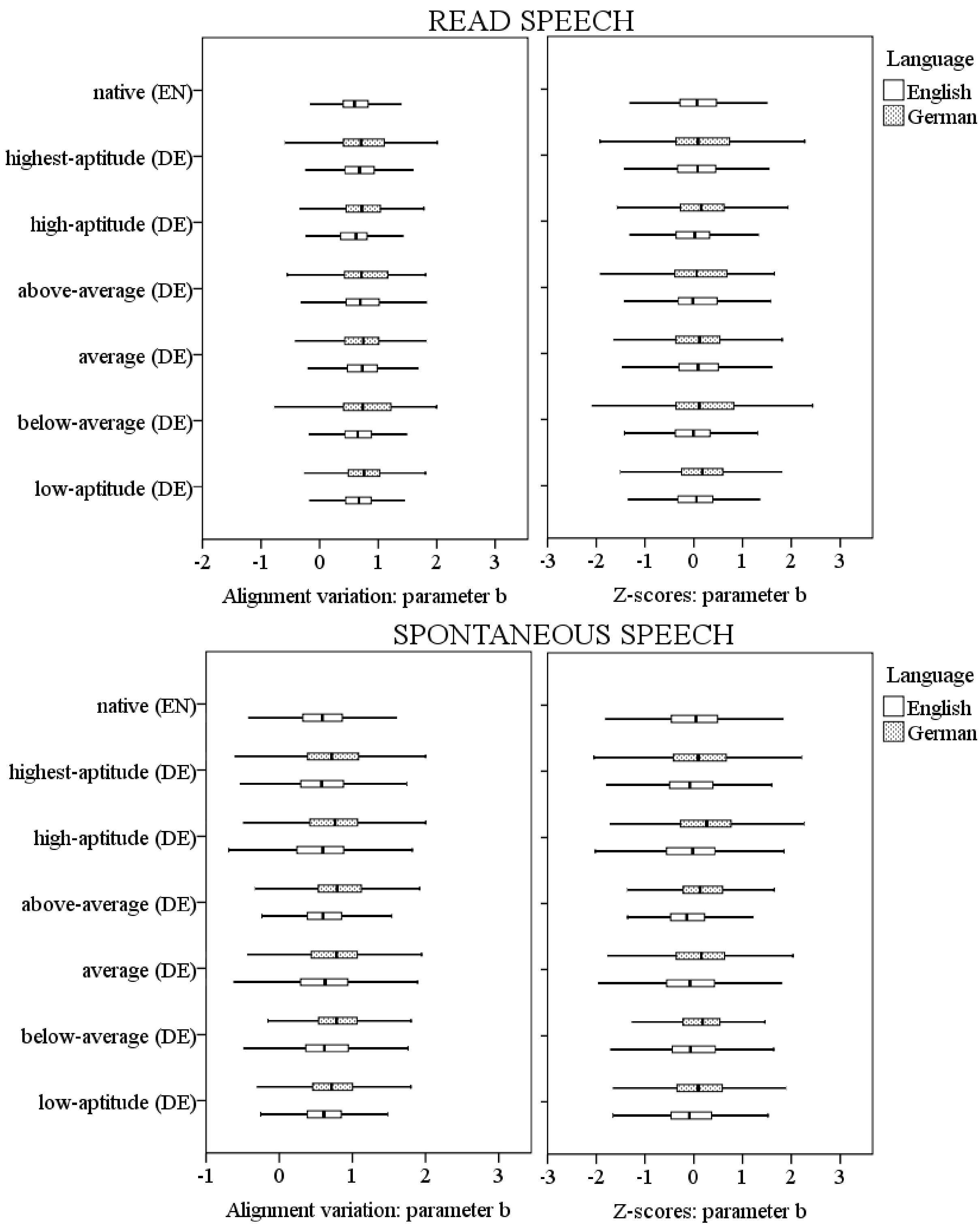


Figure 18: b parameter variation in H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German by speech type – read versus spontaneous realizations*

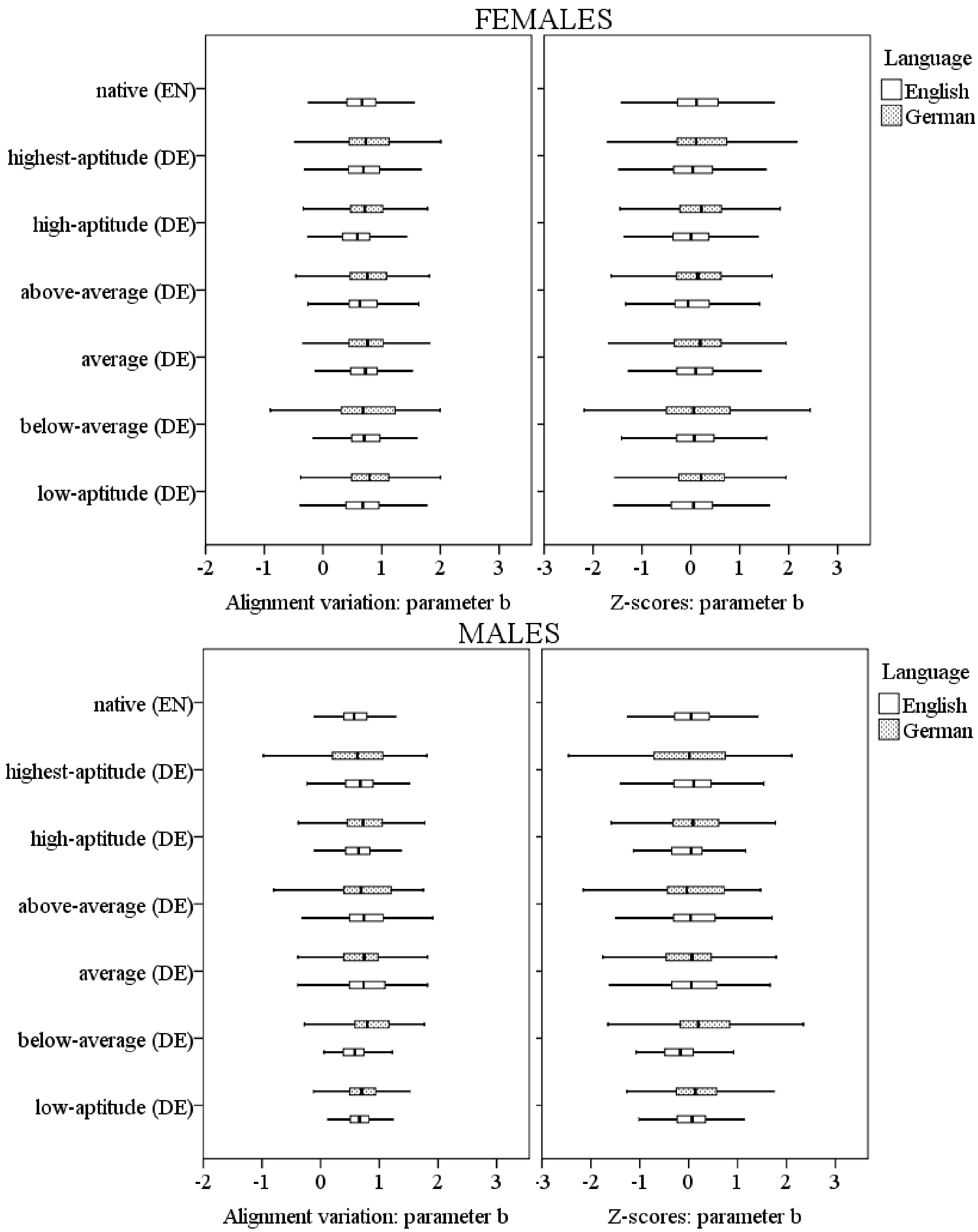


Figure 19: b parameter variation in H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German read speech by gender – male versus female realizations*

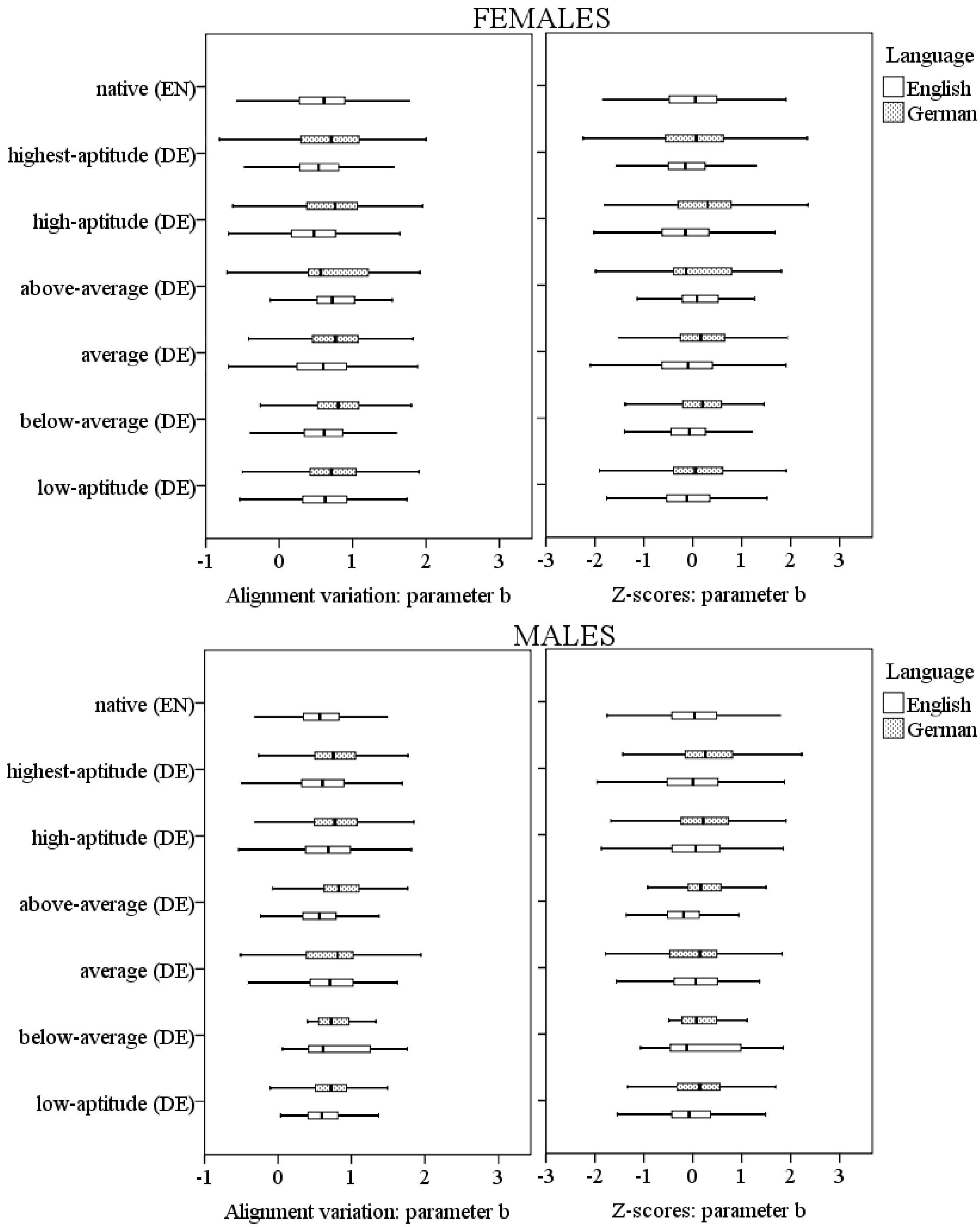


Figure 20: b parameter variation in H accents with regard to the F0 level in Hz (left charts) and the within-speaker / group variability reflected by the z-scores (right charts) in all the aptitude groups in English and German spontaneous speech by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.999					.998					.781				
	AA	.973	.997				1	1				.746	1			
	A	.998	1	.996			1	.993	1			.982	.986	.959		
	BA	.946	.994	1	.993		1	1	1	1		.404	.933	.991	.703	
	LA	.907	.991	1	.988	1	.983	.894	.995	.991	.989	.937	.999	.991	1	.835
read speech	HH															
	H	1					.986					.664				
	AA	1	1				1	1				.882	1			
	A	1	1	1			.984	1	1			.754	1	1		
	BA	.995	.999	1	.997		.934	1	.999	.998		.109	.781	.821	.612	
	LA	.759	.891	.980	.756	.991	.981	.805	.988	.733	.648	.544	1	1	.998	.836
spontaneous speech	HH															
	H	.999					1					1				
	AA	.935	.983				1	1				.962	.991			
	A	.991	1	.993			.946	.976	.995			.999	.985	.871		
	BA	.970	.996	1	.999		.920	.949	.984	.999		1	.999	.983		
	LA	1	.999	.940	.993	.973	1	1	1	.990	.970	.999	.978	.849		

Table 9: Inter-group differences in the parameter b of H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.792						.347						1					
	AA	.888	.275					.996	.583					.934	.988				
	A	.494	.010	1				.975	.018	1				.341	.604	.992			
	BA	1	.853	.976	.904			1	.319	1	1			1	.997	.925	.578		
	LA	1	.652	.960	.773	1		.999	.153	1	1	1		1	.999	.911	.316	1	
	N	.916	.998	.378	.006	.946	.799	1	.451	.985	.831	.994	.981	.947	.801	.383	.006	1	.984
read speech	HH																		
	H	.221						.142						.999					
	AA	.982	.127					1	.839					.870	.610				
	A	.936	.004	1				1	.021	1				.853	.536	1			
	BA	.994	.916	.868	.724			1	.367	1	1			.788	.939	.229	.194		
	LA	.975	.826	.751	.400	1		.999	.449	1	.981	1		.988	1	.438	.345	.977	
	N	.377	.994	.229	.004	.989	.969	.980	.558	1	.835	.994	1	.815	.990	.099	.027	.992	1
spontaneous speech	HH																		
	H	.999						.999						.920					
	AA	1	1					.960	.917					1	.926				
	A	.870	.994	1				.962	.828	.994				.764	.999	.800			
	BA	.948	.996	1	1			.994	.950	.994	1			.641	.944	.630	.990		
	LA	.764	.970	.998	1	1		.833	.642	.999	.997	.998		.991	1	.985	.991	.888	
	N	.997	1	1	.981	.993	.929	.982	.879	.991	1	1	.991	1	.846	1	.655	.595	.981

Table 10: Inter-group differences in the parameter b of H accents in the English part of the corpus*

Parameter *d*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.001					.000					.000				
	AA	.000	.000				.528	.971				.064	.011			
	A	.000	1	.000			.000	.619	.537			1	.000	.067		
	BA	.005	1	.004	1		.000	.332	.312	.966		.000	.000	.000	.000	
	LA	.000	.165	.114	.127	.549	.000	.000	.002	.001	.177	.102	.000	.991	.101	.000
read speech	HH															
	H	.737					.972					.000				
	AA	.000	.019				.283	.117				.059	.234			
	A	.001	.253	.499			.000	.000	.981			1	.000	.049		
	BA	.000	.037	.986	.819		.000	.000	.346	.248		.000	.000	.000	.000	
	LA	.000	.016	.957	.806	1	.000	.000	.093	.007	.979	.002	.109	1	.001	.000
spontaneous speech	HH															
	H	.000					.000					.000				
	AA	.000	.002				.765	.450				.856	.026			
	A	.089	.683	.000			.000	.739	.835			.890	.000	.283		
	BA	.997	.195	.000	.801		.003	.525	.981	.981		.022	.000	.002	.145	
	LA	.000	.969	.018	.201	.046	.000	.965	.218	.194	.159	.541	.000	.087	.990	.322

Table 11: Inter-group differences in the parameter d of H accents in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.000						.021						.000					
	AA	.000	.000					.421	.003					1	.001				
	A	.761	.001	.000				1	.002	.509				1	.000	1			
	BA	.970	.023	.000	1			.974	.004	.840	.994			.682	.000	.868	.728		
	LA	.795	.000	.000	.047	.396		1	.062	.384	.999	.956		.020	.882	.050	.031	.003	
	N	.000	.000	.819	.000	.000	.000	.955	.211	.135	.751	.561	.988	1	.000	.999	1	.514	.003
read speech	HH																		
	H	.013						.889						.000					
	AA	.000	.000					.000	.000					.963	.000				
	A	.337	.000	.000				.000	.000	.969				1	.000	.821			
	BA	.051	.000	.000	.842			.000	.000	1	.430			.281	.000	.059	.469		
	LA	.005	.000	.000	.661	1		.000	.000	1	.981	.921		.000	.000	.004	.000	.000	
	N	.000	.000	1	.000	.000	.000	.000	.000	.944	1	.343	.949	.980	.000	1	.779	.029	.000
spontaneous speech	HH																		
	H	.034						.477						.000					
	AA	.001	.000					1	.999					.986	.000				
	A	.000	.812	.000				.999	.147	1				.986	.000	.841			
	BA	.000	.050	.000	.463			.511	1	.999	.192			.304	.000	.164	.683		
	LA	.896	.711	.000	.030	.000		.661	.015	.979	.864	.023		.000	.177	.001	.000	.000	
	N	.816	.000	.013	.000	.000	.139	1	.175	1	1	.225	.784	1	.000	.983	.975	.261	.000

Table 12: Inter-group differences in the parameter d of H accents in the English part of the corpus*

Less frequent accents – L*, H*L, L*HL

Parameter *d*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
L*	HH															
	H	.007					.552					.008				
	AA	.000	.001				.931	1				.003	.998			
	A	.017	.993	.000			.016	.930	1			1	.000	.000		
	BA	.074	1	.011	.997		.004	.403	.959	.800		1	.058	.029	1	
	LA	.009	1	.001	.993	1	.000	.072	.890	.288	.999	.894	.310	.173	.723	.896
H*L	HH															
	H	.005					.179					.438				
	AA	.000	.000				.482	.915				1	.382			
	A	.017	1	.000			.002	.825	.991			.825	.012	.893		
	BA	.000	.234	.082	.170		.000	.079	1	.476		.161	.001	.218	.750	
	LA	.000	.019	.019	.010	1	.000	.052	1	.553	.996	1	.214	1	.754	.107
L*HL	HH															
	H	.824					.964					.198				
	AA	.108	.557				.991	1				.423	1			
	A	1	.909	.144			.441	.956	.999			.095	1	1		
	BA	.858	.341	.029	.753		.948	1	1	.999		.738	.100	.140	.067	
	LA	.000	.018	.994	.000	.000	.000	.000	.000	.000	.000	.221	1	1	1	.100

Table 13: Inter-group differences in the parameter d of less frequent accents (L, H*L, L*HL) in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
L*	HH																		
	H	1						.982						.000					
	AA	.076	.079					.618	.311					1	.000				
	A	1	.999	.144				.556	.197	.995				.924	.000	.999			
	BA	.952	.969	.034	.867			.509	.207	1	.997			1	.261	1	1		
	LA	.220	.230	.985	.388	.109		.001	.000	.971	.209	.876		.996	.000	1	1	1	
	N	.000	.000	1	.000	.000	.676	.000	.000	.961	.044	.812	1	1	.000	1	.857	1	.994
H*L	HH																		
	H	.000						.132						.000					
	AA	.000	.000					.332	.011					.565	.000				
	A	.948	.000	.000				.001	.000	1				.994	.000	.944			
	BA	.177	.000	.028	.739			.000	.000	1	.970			.772	.000	1	.985		
	LA	.018	.000	.022	.371	1		.140	.000	.995	.991	.784		.542	.000	.018	.269	.069	
	N	.000	.000	.785	.000	.114	.061	.956	.012	.702	.109	.031	.744	.918	.000	.924	1	.984	.013
L*HL	HH																		
	H	.605						1						.028					
	AA	.056	.001					.766	.710					.977	.011				
	A	1	.683	.020				.802	.783	.980				.809	.370	.412			
	BA	.996	.354	.299	.973			.013	.032	1	.264			.390	.001	.888	.048		
	LA	.015	.000	.998	.001	.291		.000	.000	.614	.000	.034		.925	.001	1	.173	.895	
	N	.157	.969	.000	.186	.074	.000	.908	.990	.370	.179	.001	.000	.998	.834	.930	1	.411	.881

Table 14: Inter-group differences in the parameter d of less frequent accents (L, H*L, L*HL) in the English part of the corpus*

Parameter *b*

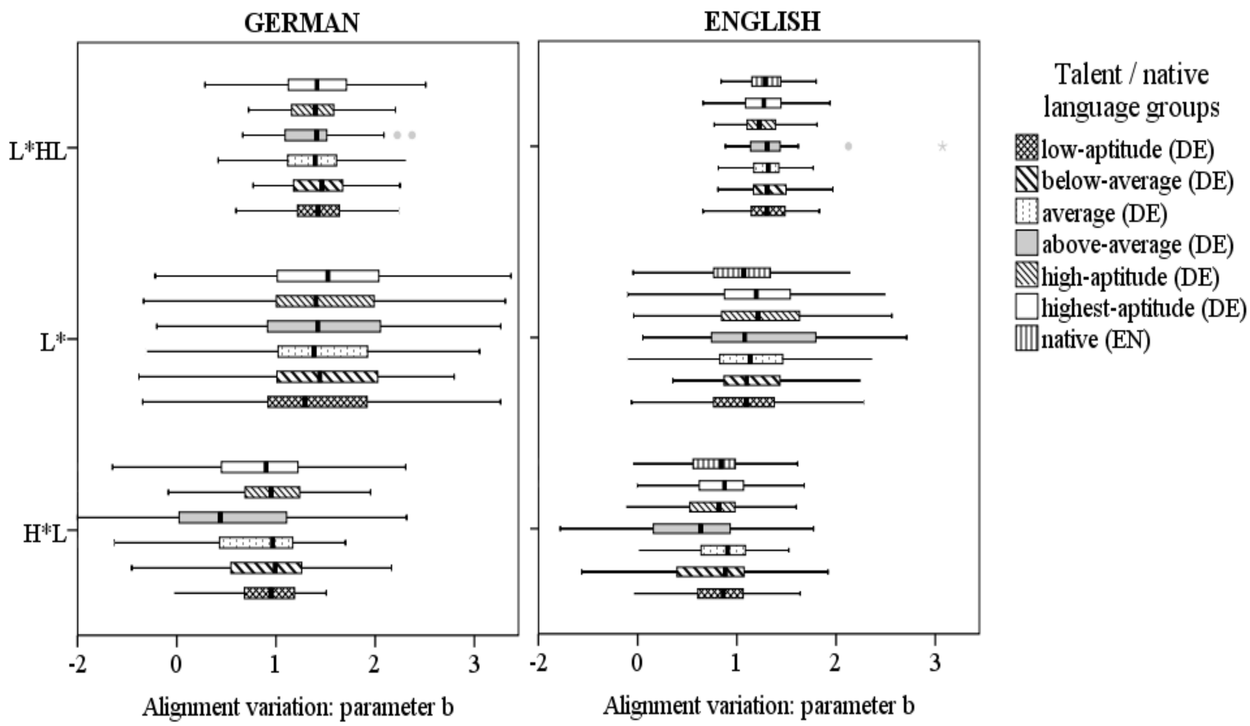


Figure 21: b parameter variation in less frequent accents (L, H*L, L*HL) in all the aptitude groups in English and German*

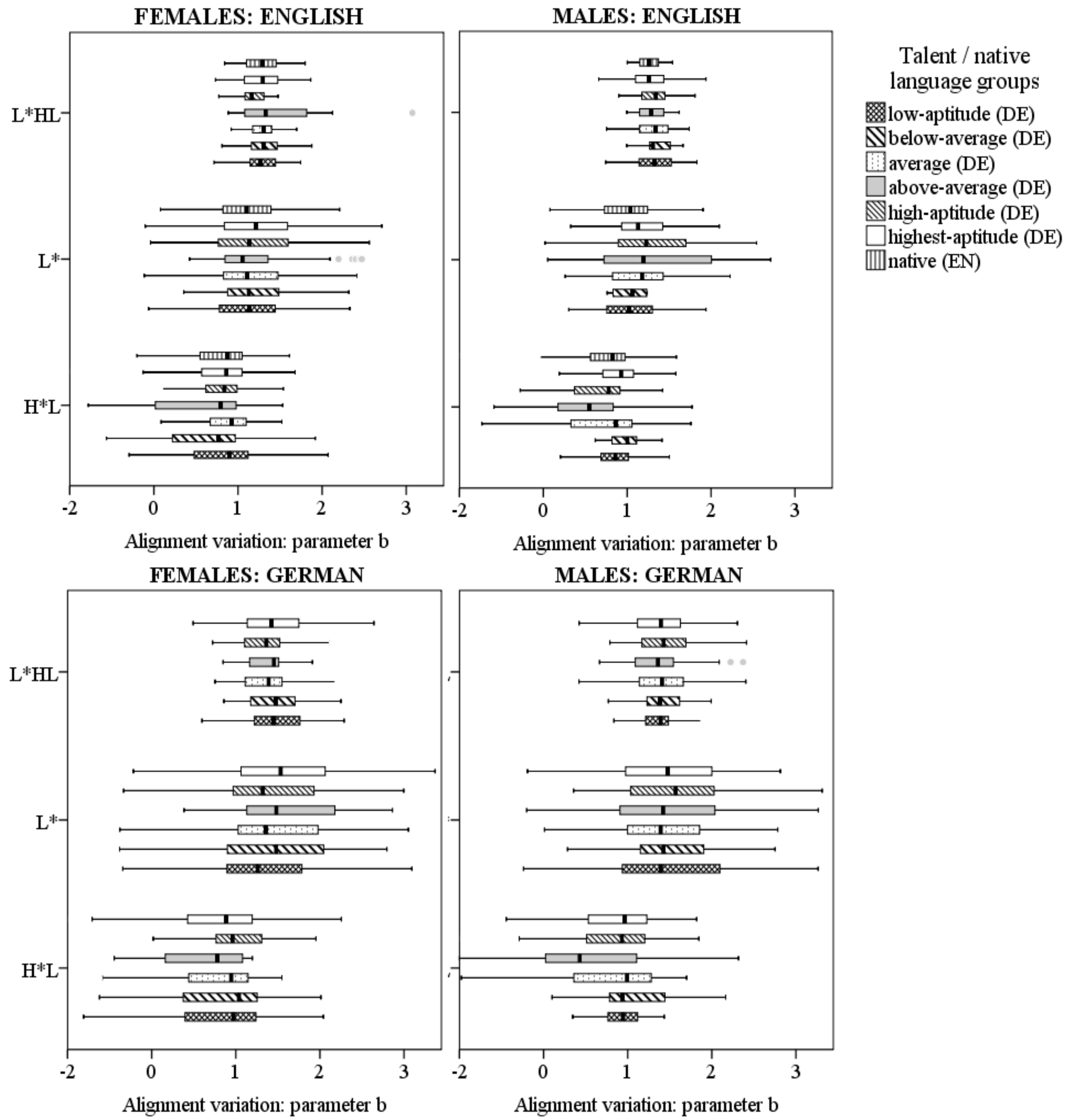


Figure 22: b parameter variation in less frequent accents (L, H*L, L*HL) in all the aptitude groups in English and German by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (p-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
L*	HH															
	H	1					.641					.768				
	AA	1	1				.999	.903				1	.801			
	A	.913	.970	.998			.970	.948	.985			.996	.300	.989		
	BA	.999	1	1	.999		.979	.996	.980	1		.997	.997	.999	.953	
	LA	.661	.794	.955	.989	.965	.256	.998	.797	.673	.940	.999	.961	1	.938	1
H*L	HH															
	H	.773					.384					1				
	AA	.038	.001				.958	.689				.144	.108			
	A	.835	.111	.309			.895	.039	.996			.991	.996	.451		
	BA	.977	1	.025	.591		1	.843	.954	.955		.937	.871	.038	.713	
	LA	1	.880	.025	.727	.992	.996	.220	.984	.997	.997	.996	.974	.012	.847	.989
L*HL	HH															
	H	.965					.410					.918				
	AA	.708	.956				.906	1				.990	.695			
	A	.789	.998	.991			.519	.999	1			1	.953	.971		
	BA	.999	.933	.681	.808		1	.575	.892	.689		1	.996	.999	1	
	LA	1	.922	.632	.699	1	.993	.196	.764	.258	1	1	.873	.996	1	1

Table 15: Inter-group differences in the parameter b of less frequent accents (L, H*L, L*HL) in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
L*	HH																		
	H	1						.994						.999					
	AA	1	1					1	1					1	1				
	A	.999	1	1				.999	1	1				1	.997	1			
	BA	.999	1	1	1			1	1	1	1			1	1	1	1		
	LA	.597	.689	.904	.828	.984		.928	1	.999	.992	.999		.814	.515	.854	.859	.999	
	N	.037	.088	.544	.138	.824	.997	.472	.975	.986	.780	.969	1	.322	.051	.571	.396	.996	1
H*L	HH																		
	H	.477						1						.025					
	AA	.001	.144					.626	.684					.001	.910				
	A	.996	.906	.010				.966	.948	.334				.121	1	.741			
	BA	.863	1	.162	.991			.492	.613	.998	.118			.997	.045	.002	.136		
	LA	.949	.991	.034	1	1		1	1	.797	.934	.835		.704	.643	.077	.902	.136	
	N	.307	1	.093	.855	1	.986	1	1	.775	.884	.781	1	.043	.951	.245	.999	.902	.952
L*HL	HH																		
	H	1						.855						.989					
	AA	.977	.928					.590	.197					1	.999				
	A	.999	.980	.997				1	.857	.527				.972	1	.999			
	BA	.994	.962	1	1			1	.789	.661	1			.953	.999	.986	.999		
	LA	.974	.903	1	.999	1		1	.679	.714	.999	1		.961	1	.995	1	1	
	N	1	1	.956	.994	.982	.955	1	.978	.506	1	.999	.995	1	.995	1	.993	.969	.985

Table 16: Inter-group differences in the parameter b of less frequent accents (L, H*L, L*HL) in the English part of the corpus*

Parameter *a2*

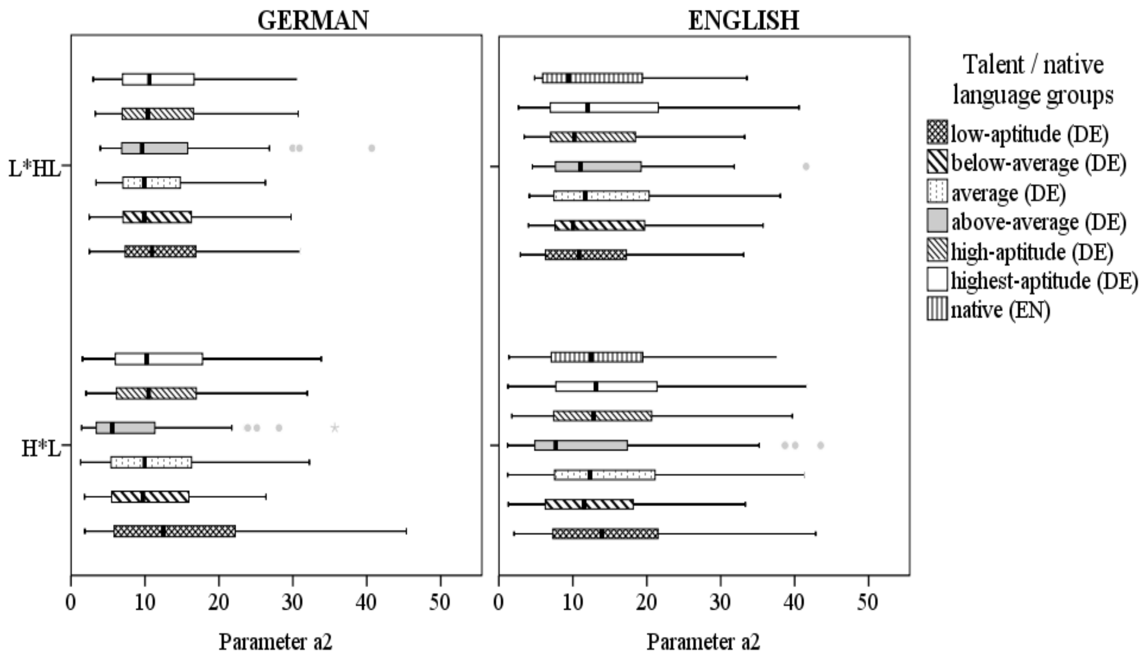


Figure 23: a2 parameter variation in less frequent accents (L, H*L, L*HL) in all the aptitude groups in English and German*

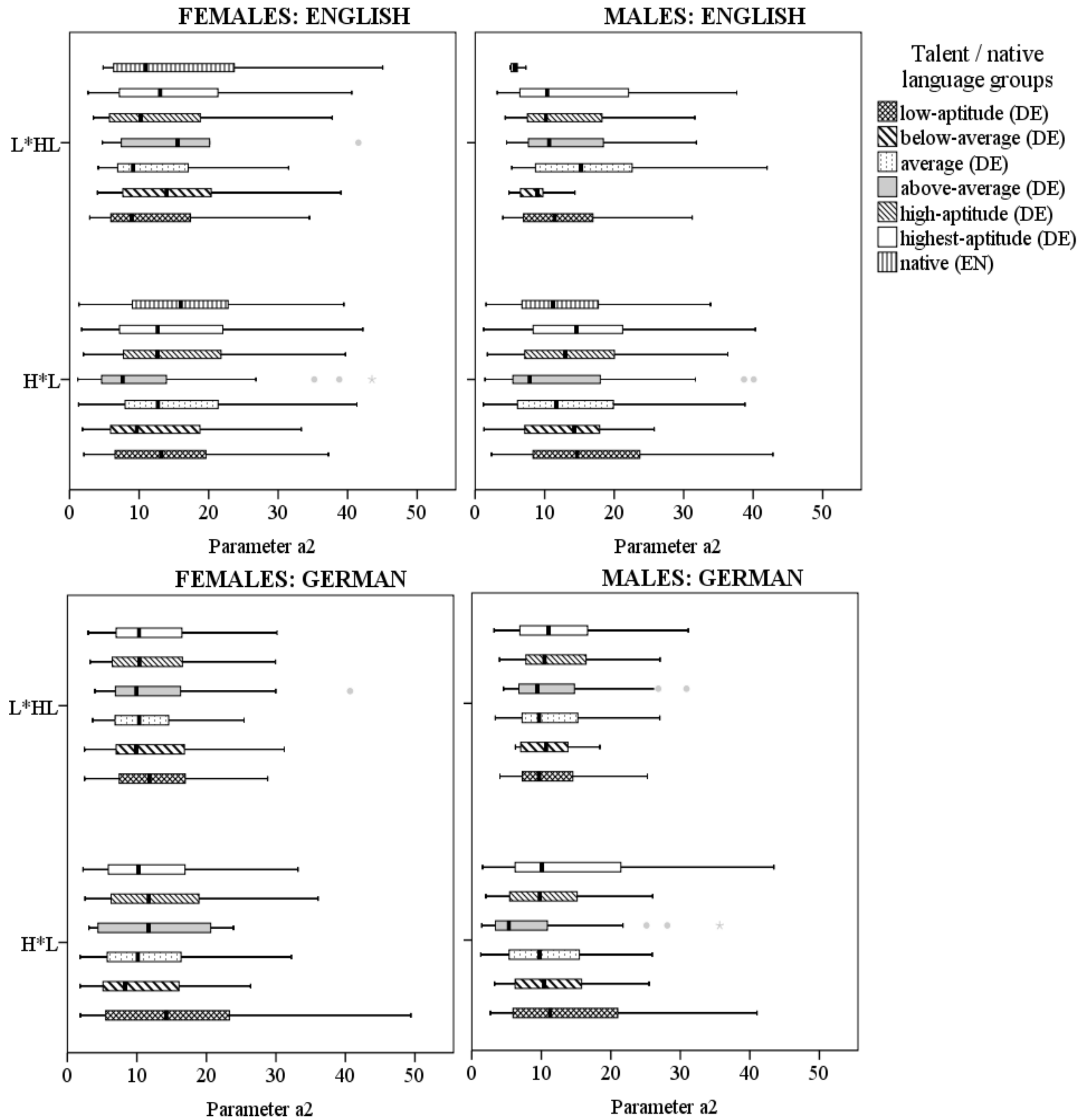


Figure 24: a2 parameter variation in less frequent accents (L, H*L, L*HL) in all the aptitude groups in English and German by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
H*L	HH															
	H	1					.594					.590				
	AA	.377	.296				1	.988				.075	.747			
	A	1	1	.351			.997	.879	1			.964	.984	.438		
	BA	.993	.981	.847	.989		1	.749	1	.991		.987	.992	.623	1	
	LA	.461	.601	.018	.561	.427	.196	.980	.947	.468	.423	1	.327	.017	.916	.974
L*HL	HH															
	H	.991					.994					.997				
	AA	.999	.977				.999	1				.963	.813			
	A	1	.934	1			.909	.999	1			.999	.948	.994		
	BA	1	.999	1	1		1	1	1	.992		.990	.961	1	.998	
	LA	.895	.997	.896	.721	.981	1	.996	1	.939	1	1	.980	.987	1	.996

*Table 17: Inter-group differences in the parameter a2 of less frequent accents (H*L, L*HL) in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
H*L	HH																		
	H	1						1						.984					
	AA	.419	.359					.830	.832					.655	.978				
	A	1	1	.448				1	1	.910				1	.995	.777			
	BA	.725	.649	.996	.754			.415	.441	1	.641			.994	1	.988	.998		
	LA	.938	.987	.111	.957	.239		.995	.995	.974	1	.918		.475	.135	.023	.604	.351	
	N	.923	.864	.828	.942	.989	.331	.898	.922	.457	.747	.060	.682	.541	.997	.999	.795	.999	.000
L*HL	HH																		
	H	.996						.949						1					
	AA	.999	1					1	.979					.999	1				
	A	1	.973	.996				.855	1	.975				.785	.754	.674			
	BA	.999	1	1	.989			1	.993	.999	.980			.878	.919	.986	.347		
	LA	.990	1	1	.946	1		.922	1	.980	1	.991		1	1	1	.836	.919	
	N	.996	1	1	.982	1	1	1	.976	1	.945	1	.969	.427	.509	.760	.065	.997	.519

*Table 18: Inter-group differences in the parameter a2 of less frequent accents (H*L, L*HL) in the English part of the corpus*

Parameter *c2*

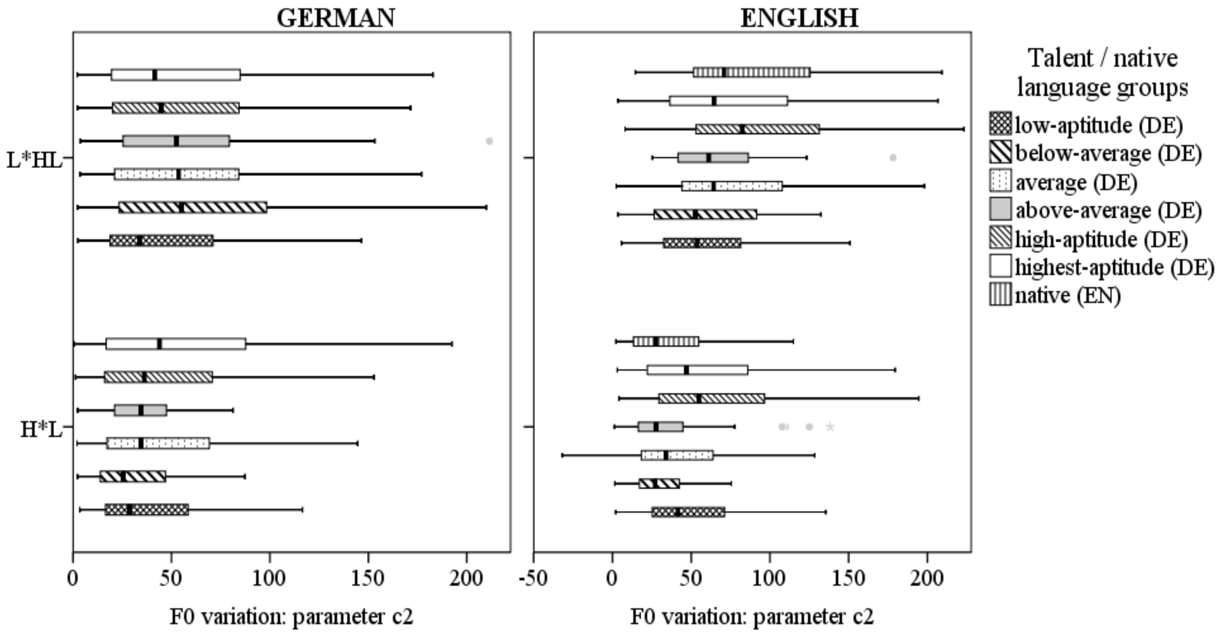


*Figure 25: c2 parameter variation in less frequent accents (H*L, L*HL) in all the aptitude groups in English and German*

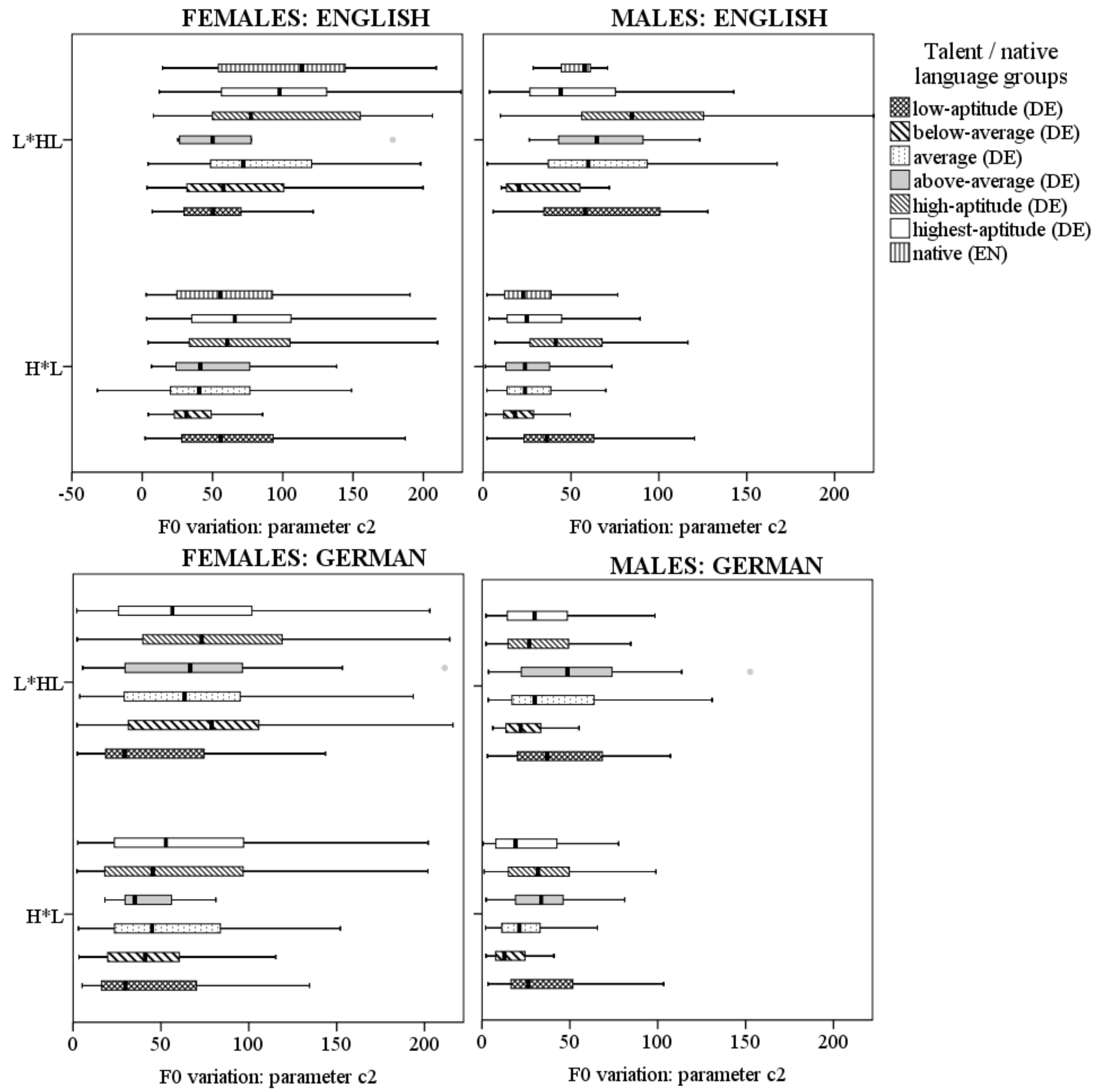


*Figure 26: c2 parameter variation in less frequent accents (H*L, L*HL) in all the aptitude groups in English and German by gender – male versus female realizations*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
H*L	HH															
	H	.866					.997					.889				
	AA	.078	.404				.843	.911				.994	.998			
	A	.386	.969	.768			.486	.843	.990			.986	.435	.827		
	BA	.021	.212	1	.582		.433	.686	1	.988		.499	.057	.227	.842	
	LA	.019	.372	.990	.888	.961	.041	.192	1	.832	1	.751	1	.983	.254	.026
L*HL	HH															
	H	.914					.613					1				
	AA	.995	1				1	.980				.436	.284			
	A	.939	1	1			1	.612	1			.866	.700	.927		
	BA	.734	.983	.984	.968		.986	.993	1	.990		.946	.963	.417	.702	
	LA	.998	.718	.960	.754	.548	.104	.001	.572	.062	.086	.497	.291	.999	.983	.505

*Table 19: Inter-group differences in the parameter c2 of less frequent accents (H*L, L*HL) in the German part of the corpus*

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
H*L	HH																		
	H	.095						1						.000					
	AA	.004	.000					.575	.563					.824	.000				
	A	.152	.000	.509				.004	.006	1				.978	.000	.998			
	BA	.000	.000	1	.323			.000	.000	.980	.660			.374	.000	.997	.898		
	LA	.846	.003	.144	.961	.047		.713	.704	.987	.872	.142		.174	.505	.014	.035	.001	
	N	.000	.000	.994	.431	.990	.38	.476	.479	.991	.869	.114	1	.895	.000	.996	1	.801	.000
L*HL	HH																		
	H	.388						1						.018					
	AA	.996	.372					.961	.940					.988	.525				
	A	1	.381	.987				.935	.915	.999				.870	.233	1			
	BA	.972	.104	1	.923			.693	.692	1	.993			.807	.010	.545	.312		
	LA	.726	.010	1	.518	.999		.046	.086	1	.316	.874		.991	.307	1	1	.540	
	N	.820	1	.677	.845	.415	.135	.998	1	.880	.719	.435	.023	1	.466	.998	.993	.946	.999

Table 20: Inter-group differences in the parameter c2 of less frequent accents (L, H*L, L*HL) in the English part of the corpus*

Boundary tones

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.981					.164					.010				
	AA	.000	.000				1	.813				.941	.350			
	A	.457	.926	.001			.572	.922	.947			1	.001	.918		
	BA	.999	1	.000	.901		.220	1	.795	.918		.040	.000	.006	.019	
	LA	.354	.840	.005	1	.820	.001	.792	.480	.093	.923	1	.004	.909	1	.039
read speech	HH															
	H	.915					1					.050				
	AA	.738	.293				.650	.679				.528	.986			
	A	1	.919	.597			1	1	.675			.994	.058	.687		
	BA	.705	.997	.163	.679		.940	.908	.386	.776		.502	.002	.039	.215	
	LA	.647	.999	.137	.556	1	.958	.928	.409	.768	1	.426	.924	1	.570	.023
spontaneous speech	HH															
	H	.527					.034					.432				
	AA	.000	.002				.980	.998				1	.350			
	A	.283	1	.002			.318	.874	1			.998	.094	1		
	BA	.749	1	.023	1		.759	.964	1	1		.170	.002	.289	.247	
	LA	.001	.258	.202	.274	.694	.000	.869	.948	.161	.535	.550	.002	.785	.721	.814

Table 21: Inter-group differences in parameter d of default intonation phrase boundary tones (%) in the German part of the corpus

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.080						.919						.000					
	AA	.007	.000					.669	.232					.890	.000				
	A	.939	.456	.000				.943	.249	.931				.638	.000	1			
	BA	.129	1	.000	.519			.988	1	.373	.566			.601	.000	.144	.052		
	LA	1	.064	.008	.917	.109		.349	.019	1	.871	.097		.040	.000	.824	.791	.002	
	N	.162	.000	.704	.001	.000	.178	.117	.004	1	.513	.025	.997	.977	.000	.999	.983	.209	.273
read speech	HH																		
	H	.742						.998						.000					
	AA	.006	.000					.853	.658					1	.000				
	A	.945	.092	.059				.888	.555	.996				.990	.000	1			
	BA	.996	.470	.130	1			.368	.140	1	.929			.882	.000	.716	.496		
	LA	.892	.057	.085	1	1		.035	.005	1	.493	1		.658	.001	.928	.948	.135	
	N	.146	.001	.789	.642	.765	.746	.902	.636	.999	1	.985	.835	.987	.000	1	1	.480	.969
spontaneous speech	HH																		
	H	.343						.965						.000					
	AA	.740	.036					.940	.667					.905	.009				
	A	.119	1	.012				1	.945	.919				.757	.000	1			
	BA	.000	.198	.000	.107			.022	.375	.037	.005			.696	.000	.270	.171		
	LA	.996	.862	.480	.729	.006		1	.976	.955	1	.055		.301	.043	.993	.982	.053	
	N	.975	.037	.972	.003	.000	.786	.506	.086	1	.275	.000	.676	1	.000	.954	.863	.596	.405

Table 22: Inter-group differences in parameter d of default intonation phrase boundary tones (%) in the English part of the corpus

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.950					.920					.000				
	AA	.000	.000				1	.995				.000	.382			
	A	1	.971	.000			.291	.025	.978			.028	.281	.001		
	BA	.786	.362	.000	.689		.992	.754	1	.945		1	.025	.001	.413	
	LA	.022	.234	.119	.020	.004	.000	.000	.236	.003	.005	.020	.460	.005	1	.358
read speech	HH															
	H	.973					.829					.000				
	AA	.000	.000				.997	.907				.000	.990			
	A	.975	.638	.000			.315	.015	.999			.052	.121	.509		
	BA	.991	1	.000	.852		.886	.320	1	.998		.993	.002	.012	.201	
	LA	.032	.003	.185	.181	.051	.000	.000	.384	.002	.012	.000	.912	.999	.692	.020
spontaneous speech	HH															
	H	.427					1					.201				
	AA	.117	.895				.998	.999				.000	.022			
	A	.996	.080	.019			.907	.851	.965			.658	.992	.016		
	BA	.716	.066	.015	.864		.999	1	1	.868		.994	.994	.436	1	
	LA	.810	.997	.744	.408	.189	.319	.268	.804	.819	.403	1	.232	.000	.681	.995

Table 23: Inter-group differences in parameter d of high intonation phrase boundary tones (H-H%) in the German part of the corpus

		STATISTICAL SIGNIFICANCE (p-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.090						.656						.000					
	AA	.000	.000					.295	.017					.879	.000				
	A	.997	.190	.000				.488	.003	.885				1	.000	.916			
	BA	.030	.984	.000	.068			.528	.013	.972	1			.894	.000	.540	.848		
	LA	1	.103	.000	1	.036		.000	.000	1	.010	.341		.017	.000	.537	.014	.079	
	N	.435	.000	.013	.048	.000	.218	.679	.018	.894	1	1	.049	1	.000	.849	1	.884	.008
read speech	HH																		
	H	.950						.401						.012					
	AA	.000	.000					.272	.005					.924	.000				
	A	.171	.007	.000				.027	.000	.998				.916	.000	1			
	BA	1	.978	.000	.427			.074	.000	1	1			.706	.002	.973	.960		
	LA	.577	.070	.000	.992	.783		.000	.000	1	.520	.943		.832	.261	.086	.044	.166	
	N	.001	.000	.152	.389	.011	.125	.007	.000	1	.930	.995	1	.972	.000	1	1	.951	.170
spontaneous speech	HH																		
	H	.045						1						.000					
	AA	.999	.098					.981	.945					.001	.035				
	A	.005	1	.034				1	1	.971				.557	.000	.121			
	BA	.000	.125	.000	.164			1	1	.975	1			---	---	---	---		
	LA	.371	.989	.380	.921	.027		.461	.229	1	.181	.502		.021	.001	.966	.602	---	
	N	.972	.225	.901	.040	.000	.815	1	.999	.985	1	1	.350	.928	.000	.006	.956	---	.125

Table 24: Inter-group differences in parameter d of high intonation phrase boundary tones (H-H%) in the English part of the corpus

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	1					.961					.646				
	AA	.344	.412				.566	.872				.998	.976			
	A	1	1	.271			.738	.998	.945			.832	1	.997		
	BA	.887	.918	.988	.837		.636	.928	1	.979		.934	.388	.850	.522	
	LA	.091	.143	1	.052	.981	.271	.722	1	.864	1	.834	.999	.997	1	.525
read speech	HH															
	H	.987					1					.772				
	AA	.332	.145				.720	.754				.997	.977			
	A	1	.972	.334			.931	.959	.941			.822	1	.994		
	BA	.678	.388	.997	.690		.458	.518	1	.835		.960	.438	.861	.474	
	LA	.193	.061	1	.181	.999	.405	.468	1	.801	1	.804	1	.992	1	.458
spontaneous speech	HH															
	H	.983					.787					.800				
	AA	.977	1				.938	1				1	.905			
	A	1	.968	.966			.747	1	1			1	.783	1		
	BA	1	.998	.993	1		1	1	.999	1		---	---	---		
	LA	.740	.973	1	.689	.969	.894	1	1	1	1	1	.881	.999	.998	

Table 25: Inter-group differences in parameter d of low intermediate phrase boundary tones (L-) in the German part of the corpus

		STATISTICAL SIGNIFICANCE (p-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	1						1						.486					
	AA	.360	.367					.940	.923					.991	.307				
	A	.975	.992	.106				1	.999	.978				.695	1	.451			
	BA	1	1	.490	.996			1	1	.966	1			.945	.187	1	.298		
	LA	.988	.985	.928	.817	.992		.998	.995	.999	1	.999		1	.870	.992	.934	.959	
	N	.023	.048	1	.001	.214	.915	1	1	.964	1	1	1	.757	.002	1	.022	1	.919
read speech	HH																		
	H	1						.998						.984					
	AA	.763	.907					.952	.872					.999	.934				
	A	.999	1	.904				1	.979	.986				1	.999	.989			
	BA	.360	.668	1	.632			.984	.918	1	.999			.447	.090	.966	.236		
	LA	.958	.996	.998	.996	.990		.998	.972	.999	1	1		.999	.928	1	.990	.922	
	N	.001	.030	.997	.011	.990	.716	1	1	.924	.996	.966	.991	.424	.029	.995	.142	.996	.981
spontaneous speech	HH																		
	H	.992						1						.325					
	AA	.926	.686					1	1					1	.457				
	A	.393	.894	.131				1	1	.999				.993	.602	.997			
	BA	.031	.226	.010	.864			.810	.915	.933	.908			---	---	---	---		
	LA	1	.995	.992	.717	.175		1	1	1	.999	.903		.955	.987	.956	.903	---	
	N	1	.981	.902	.228	.012	1	.980	.997	.991	.998	.986	.987	1	.137	1	.995	---	.912

Table 26: Inter-group differences in parameter d of low intermediate phrase boundary tones (L-) in the English part of the corpus

		STATISTICAL SIGNIFICANCE (p-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
all corpus	HH															
	H	.011					.214					.026				
	AA	.000	.000				.805	1				1	.049			
	A	.011	1	.000			.002	.893	.998			1	.037	1		
	BA	.422	.996	.000	.999		.012	.747	.980	.992		.733	.005	.733	.559	
	LA	.000	.099	.008	.044	.146	.000	.024	.688	.242	.903	1	.008	1	.999	.709
read speech	HH															
	H	.369					1					.273				
	AA	.000	.005				.971	.989				.999	.587			
	A	.320	1	.003			.603	.825	1			.995	.464	1		
	BA	.821	1	.011	1		.511	.694	.999	.996		.781	.034	.609	.481	
	LA	.000	.257	.432	.160	.332	.025	.091	.939	.594	.972	.987	.563	1	1	.417
spontaneous speech	HH															
	H	.019					.000					.048				
	AA	.000	.000				.837	1				.936	.003			
	A	.015	1	.000			.000	.995	1			.874	.001	1		
	BA	.447	1	.003	1		.000	.788	.984	.930		.981	.282	1	1	
	LA	.000	.704	.001	.772	.965	.000	.576	.991	.872	1	.180	.000	.851	.877	.998

Table 27: Inter-group differences in parameter d of low intonation phrase boundary tones (L-L%) in the German part of the corpus

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
all corpus	HH																		
	H	.000						.440						.000					
	AA	.000	.000					.012	.000					.805	.000				
	A	1	.000	.000				.020	.000	.593				.850	.000	.165			
	BA	.925	.355	.000	.806			.373	.003	.500	1			.978	.000	1	.648		
	LA	.039	.000	.000	.115	.010		.009	.000	.831	.997	.977		.997	.000	.497	.994	.887	
	N	.000	.000	.377	.000	.000	.007	.646	.004	.166	.768	.992	.493	.230	.000	1	.002	1	.044
read speech	HH																		
	H	.029						.475						.000					
	AA	.000	.000					.010	.000					.776	.000				
	A	.426	.000	.000				.011	.000	.686				.973	.000	.259			
	BA	.940	.007	.000	1			.087	.000	.795	1			.878	.000	1	.489		
	LA	.001	.000	.014	.421	.462		.008	.000	.869	.999	1		1	.000	.935	.841	.965	
	N	.000	.000	.837	.000	.000	.026	.661	.008	.182	.722	.847	.525	.221	.000	1	.005	1	.558
spontaneous speech	HH																		
	H	.102						.999						.000					
	AA	.150	.000					.996	.986					.920	.000				
	A	.026	1	.000				1	1	.994				.954	.000	1			
	BA	.000	.059	.000	.054			.747	.954	.889	.794			---	---	---	---		
	LA	.974	.661	.037	.461	.000		1	.991	.999	.999	.633		.031	.519	.017	.015	---	
	N	.998	.008	.235	.001	.000	.749	.998	1	.984	1	.915	.980	.840	.000	1	1	---	.000

Table 28: Inter-group differences in parameter d of low intonation phrase boundary tones (L-L%) in the English part of the corpus

		STATISTICAL SIGNIFICANCE (*p*-values, Scheffé post-hoc tests)														
		all speakers					females					males				
		HH	H	AA	A	BA	HH	H	AA	A	BA	HH	H	AA	A	BA
North Wind and the Sun	HH															
	H	.000					.000					.984				
	AA	1	.001				-----	-----				-----	-----			
	A	.000	.976	.000			.000	.062	----			.020	.005	-----		
	BA	.000	.059	.000	.084		-----	-----	-----	-----		-----	-----	-----	-----	
	LA	.017	.000	.316	.000	.000	.020	.000	-----	.008	-----	1	.951	-----	.000	-----
Mrs. McWilliams	HH															
	H	.002					.796					.877				
	AA	.000	.282				-----	-----				-----	-----			
	A	.228	.044	.002			.560	.999	-----			.153	.000	-----		
	BA	.123	.997	.243	.818		-----	-----	-----	-----		-----	-----	-----	-----	
	LA	.000	.989	.380	.000	.937	.055	.445	-----	.056	-----	1	.669	-----	.000	-----

Table 29: Inter-group differences in intonational text interpretation in the German part of the corpus

		STATISTICAL SIGNIFICANCE (p-values, Scheffé post-hoc tests)																	
		all speakers						females						males					
		HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA	HH	H	AA	A	BA	LA
North Wind and the Sun	HH																		
	H	.979						.991						.932					
	AA	.521	.839					------	-----					------	-----				
	A	.000	.000	.963				.000	.001	-----				.640	.973	-----			
	BA	.011	.061	.983	1			------	-----	-----	-----			------	-----	-----	-----		
	LA	.052	.000	.001	.000	.000		.281	.045	-----	.000	-----		.088	.000	-----	.000	-----	
	N	.000	.000	.816	.618	1	.000	.178	.403	-----	.000	-----	.000	.003	.002	-----	.000	-----	.000
Mrs. McWilliams	HH																		
	H	.656						.116						1					
	AA	.443	968					------	-----					------	-----				
	A	.000	.002	.977				.000	.459	-----				.283	.027	-----			
	BA	.001	.080	.927	.996			------	-----	-----	-----			------	-----	-----	-----		
	LA	.573	.000	.014	.000	.000		.108	.000	-----	.000	-----		.999	1	-----	.000	-----	
	N	.000	.000	.017	.000	.239	.000	.000	.116	-----	.609	-----	.000	.000	.000	-----	.000	-----	.000

Table 30: Inter-group differences in intonational text interpretation in the English part of the corpus