3 Phonetic Realizations of Phonological Intonation Categories in German

Claudia K. Ohl

Abstract

This study investigates the influence of varying segmental structures on the realizations of utterance-final rising and falling intonation contours. Following Grabe's study on adjustment strategies in German, i.e. truncation and compression, an experiment similar to hers was carried out, using materials with decreasing stretches of voicing in questions, lists, and statements. However, the obtained results could not confirm the idea of such common adjustment strategies. Instead, strong variations were found as to how the phrase-final intonation contours were modified to match the respective amounts of voicing: the strategies varied strongly accross the different word groups and from speaker to speaker.

3.1 Introduction

Truncation and compression are assumed to be the two basic mechanisms for intonation contour modification when the duration of voicing decreases, either because only few voiced phones are available or because of a high speech rate.

How do intonation contours adapt to segmental strings of decreasing length? By compressing the intonational patterns, by truncating them or by a combination of both mechanisms? Some of the studies addressing these questions in a number of languages are summarized in the following paragraphs.

3.1.1 Adjustment strategies of intonation contours

Much of the literature on German intonation, so far, has been focussed on establishing a phonological inventory of intonation categories. Since the acoustic correlate of intonation, the fundamental frequency (F0), occurs only in voiced segments, the question of what happens to intonation in utterances with only little voiced material has been raised and addressed in several studies. Féry found that the alignment of melodies to texts depends largely on the utterances' segmental structure (Féry 1993, p. 54).

According to von Essen (1964), neither an extension nor a shortening cause a change to the underlying melody of an utterance (von Essen 1964, p. 21). This means that the duration of the material available has no influence on the perception of intonation, i.e. phonologically similar contours are perceived as equal even if the utterances' durations may vary gradually (see Fig. 3.1).

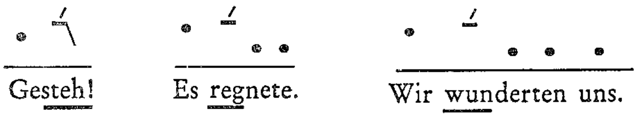

Fig. 3.1: Notation of intonation contours according to von Essen (1964, p. 21).

When the F0 movement has to be carried out within a single syllable (e.g. <*Mann*>), or even a single accentuated vowel (e.g. <*Fisch*>), then the resulting contour is *"an automatic phonetic consequence that can not and may not be included in the representation of intonation"* (Bierwisch 1966, p. 142).

Niebuhr & Ambrazaitis (2006) found a strong correlation between the duration of the slope movement and segmental chains. They concluded that the slope's shape is determined by the syllabic structure, which means that it is not the segmental structure that is matched to an intended intonation contour but rather the course of F0 that is matched to the segments' duration (varying, however, from sound to sound and speaker to speaker).

A similar result was found by Dilley et al. (2005) in their study on alignment effects of F0 peaks and valleys relative to the neighboring vowel and syllable boundary position. They found that the components of a tone sequence are not connected to each other but to the segmental chain and by that possess an independent timing.

Further proof for this connection is delivered by the *Segmental Anchoring Hypothesis* proposed by Atterer & Ladd (2004), claiming that tones are tied to specific positions (or *anchors*) in the segmental chain.

Truncation and compression

Generally, two strategies for intonation contour modifications are assumed: *truncation* and *compression*. The first to mention these strategies were Erikson & Alstermark (1972) in their study of adjustment strategies of intonation to varying vowel durations in Swedish. *Truncation* holds the idea that a contour is cut off during its course, opposing the principle of *rate adjustment* where the complete F0 movement is carried out during the given segment duration (Fig. 3.2).

At decreasing vowel duration, contours are modified along the time axis and, additionally, either adjusted by faster movements or cut off along the frequency axis.

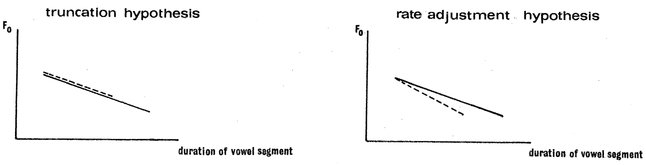

Fig. 3.2: Schematic representation of phonetic adjustment strategies of intonation categories (dashed line: modified contour) (Erikson & Alstermark 1972, p. 15).

Bannert & Bredvad-Jensen (1975) later conducted another study on dialect-based accent realizations in Swedish and introduced the term *compression* for *rate adjustment*.

Gartenberg & Panzlaff-Reuter (1991) analyzed intonation contours in German utterances with short stretches of voicing in the accented vowel such as *Sie strickt* (engl.: *She knits*).

Ladd (1996) also included a chapter on *'truncation' vs. 'compression'* in his book *Intonational Phonology*, analyzing these two effects in phrase-final accented syllables for several languages, among them German. When several tones have to be realized in a monosyllabic utterance, German behaves differently than English, where all tones are realized within the given syllable and are, thereby, compressed (Fig. 3.3).

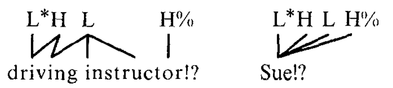

Fig. 3.3: L*H L H% on a poly- and a monosyllabic English utterance (Ladd 1996, p. 132).

While the contour stretches to five syllables in one case (<*driving instructor?!*>), the exact same contour is also possible on a monosyllabic word (<*Sue?!*>). German, on the other hand avoids compression and would change the type of pitch accent to one that can be realized more easily within a single syllable. Examples can be seen in Figure 3.4.

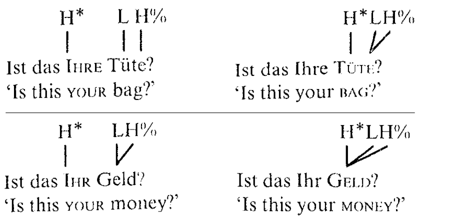

Fig. 3.4: H* L H% realizations in German on words with varying amounts of syllables (Ladd 1996, pp. 133,134).

While in English four tones can be assigned to one syllable quite easily, the tone-text-association suggested in the right column of the German examples would lead to unnaturally sounding utterances, while at the same time spreading of the tones to the last but one word of the utterance is possible. German examples for truncated intonation contours are presented by the contours used when answering the phone by giving the last name. Normally, this is done by using a

Fig. 3.5: Pitch accent realizations used in German when answering the phone in words with varying numbers of syllables (Ladd 1996, pp. 134,135).

low rise (L* H). If, instead a high rise (H* H) is used, it supposedly conveys unfriendliness and boredom (see Fig. 3.5; Ladd 1996, pp. 134f.).

When producing these contours on a monosyllabic word, the contours realized are, according to Ladd (1996), not the same high rise that would convey boredom in other cases. Much rather, it is a low rise, truncated on the left side (Fig. 3.5, right). However, Ladd himself limits the validity of his observations but at the same time concludes that German tends to avoid compression, showing strong tendencies towards truncation.

While investigating a matter quite similar in Dutch, i.e. phrase-final accent adjustment effects during little voiced material, Hanssen et al. (2007) found contour-dependent choices of compression and truncation. Simple rising and falling contours were compressed and slighty truncated. The rise-fall studied was, in addition, reduced in pitch range, implying compression. Furthermore, a reduction of absolute change of F0 could be observed, implying again truncation. So, compression along the time axis can be accompanied by compression along the frequency axis, leading the authors to the conclusion that truncation and compression are not exclusive categories.

However, adjustment strategies do not vary only from language to language, but may also differ among varieties of one language, as was already stated for Swedish by Bannert & Bredvad-Jensen (1975).

Peters & Pfitzinger (2008) conducted a study on the perception of phrase-final intonation contours, focussing, in particular, on the influence of duration and F0 interval variations in final vowels (from the utterance *"Ulf stickt"*) on the perception of interrogative and non-interrogative utterances. Figure 3.6 summarizes the stimuli used in the perception experiment. The voiced segment varied with respect to slope and duration between 20 ms and 50 ms. F0 intervals were between -4 st and +6.5 st, in steps of 1.5 st. In the

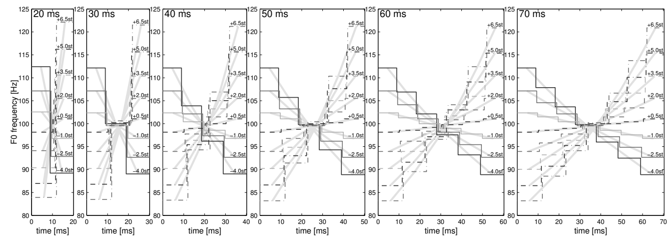

Fig. 3.6: Interpolated F0 contours (light gray lines) for the final voiced segment of the utterance *"Ulf stickt."* Additionally, contours are shown in discrete semitone steps (black) (x-axis: duration in ms; y-axis: F0 in Hz) (Peters & Pfitzinger 2008).

experiment, test subjects had to assign the heard utterances to sentence types, i.e. statement or question.

The authors found that a voicing duration of 50 ms and an interval of two semitones is enough to produce communicatively relevant intonation contours, providing some idications that the slope of a contour is not the most important trigger for perception. Isačenko & Schädlich (1966, p. 58) were the first to show that even an F0 interval of one semitone is sufficient to yield a perceivable prominence of a syllable, and they assumed that the actual F0 interval is used by the speaker for *"the expressive and emotional connotations which can be conveyed by intonation"* (Isačenko & Schädlich 1970, p. 57).

Uhmann (1997) reports about two phenomena called "Quetschung" (squeezing) and "Dehnung" (stretching) in the course of her discussion on tone-text-association in German (see Fig. 3.7).

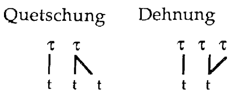

Fig. 3.7: Schematic representation of the two adjustment strategies "Quetschung" (squeezing) and "Dehnung" (stretching) that occur at a given uneven amount of tones and tone bearing units (Uhmann 1997, p. 181).

Squeezing occurs, when more tones than tone bearing units are available since all tones must be assigned. Stretching, on the other hand, occurs in the reverse case of tone bearing units exceeding the number of tones (Uhmann 1997, pp. 180,181).

```
Fe*rn seh er    Freu*nd
  | \ | /         | \
  H*+T           H*+T
```

Fig. 3.8: Squeezing and stretching in the example of *"Fernseher"* (*television*) and *"Freund"* (*friend*) (Uhmann 1997, p. 182).

Consequently, the realization of a contour on the two words has to differ because the tones have different limitations along the time axis. Concerning truncation and compression, it could be said that they are phonetic realizations of the squeezing phenomenon since it describes the adjustment of an intonation contour to a lack of tone bearing units.

Now following is a summary of Grabe's contrastive studies on English and German intonation. She found that while English is a purely compressing language, German only compresses rises and truncates falls. Since that contrastive element is of no relevance for the present paper, the main focus will be on the results concerning German intonation.

Grabe, 1998

First of all, Grabe (1998a,b) established a tone inventory for German and English. Following that, she conducted two experiments designed specifically to investigate truncation and compression, which will be introduced in this section.

The first experiment was focussed on the realizations of H*+L in words with a decreasing amount of voiced segments. The second experiment was intended to analyze the adjustment of H*+L in combination with the two boundary tones H% (rising) and 0% (unspecified). While the first experiment was designed to study both strategies (truncation and compression), the second was limited to analyze compression in German only.

Experiment I

In order to gain phrase-final realizations of intonation categories on words with decreasing durations of voiced segments, Grabe used three target words per language. Those were last names with gradually decreasing amounts of sonorant segments, embedded in a dialogue about a lottery winner.

German:	*<Schiff>*	*<Schief>*	*<Schiefer>*
English:	*<Schift>*	*<Scheaf>*	*<Scheafer>*

All words occured in phrase-final position of a statement as well as a question, triggering falling contours in statements and rising contours in questions. The test subjects were 12 female Germans from Braunschweig between the age of 16 and 18 and 12 female native English speakers from South-England between the age of 21 and 24. The test subjects were told that they were participating in a recording for intonation excercises for lerners of the respective language. Since it would be best for learners to hear natural intonation, test subjects were asked to speak, i.e. read, naturally and produce all sentences the same.

Grabe's hypotheses stated that English would compress rising and falling contours, whereas German would truncate falls and compress rises.

The results of the first experiment are summarized in Figures 3.9 and 3.10 below.

As they show, *<Schiff>*/*<Shift>* have the shortest durations, followed by *<Schief>*/*<Sheaf>*. Clearly the longest are *<Schiefer>*/*<Sheafer>*.

As far as the F0 contour adjustment to these duration relations is concerned, both languages behave similarly for rises,

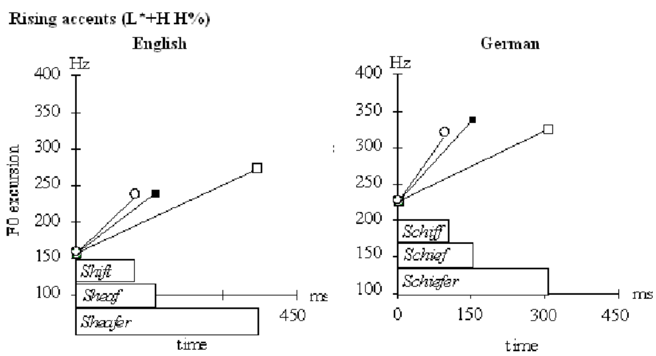

Fig. 3.9: Summarized results from the first experiment for the realizations of rising contours (x-axis: duration in ms; y-axis: F0 in Hz; rate of F0 change in Hz/ms) (Grabe 1998a, ch. 5.3.2).

i.e. they compress. This can be seen from the degree of the slope that results from an increase in the rate of F0 change. While the words with the shortest voiced segment (<*Schiff*> for German and <*Shift*> for English) have the greatest rate of change, that rate decreases with increasing duration, although they all end on a similar frequency level.

For falling contours, Grabe concluded with the following results:

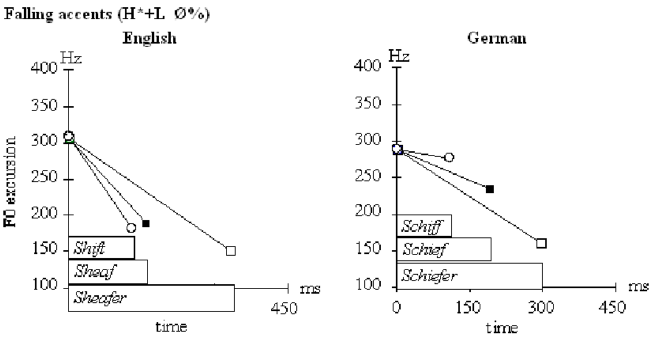

Fig. 3.10: Summarized results from the first experiment for the realizations of falling contours (x-axis: duration in ms; y-axis: F0 in Hz; rate of F0 change in Hz/ms) (Grabe 1998a, ch. 5.3.2).

English and German display different strategies to realize falling contours in a short time span. English modifies falls just like rises by compressing them. The rate of F0 change, therefore, increases with decreasing word length. German, on the other hand, does not level the duration available by increasing the rate of change, but decreases it instead.

The course of the contour is (at least theoretically) gradually shortened when duration decreases. A representation of such shortening is given in Figure 3.11.

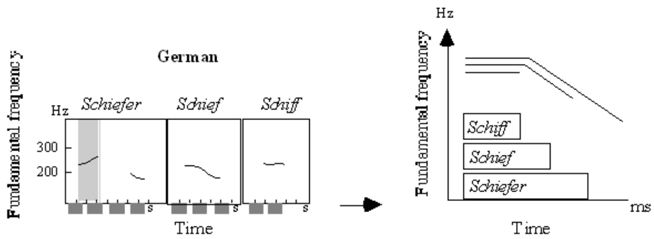

Fig. 3.11: Example for gradual shortening of a contour at decreasing word duration by truncation. On the left side the actual F0 contour and on the right side, a stylization of the adjustment (x-axis: duration in s (left) and ms (right); y-axis: F0 in Hz) (Grabe 1998a, ch. 5.3.3.2).

The change at the transition of <*Schiefer*> to <*Schief*> is still gradual, but loses that quality at the transition from <*Schief*> to <*Schiff*>. According to Grabe (1998a,b), the contour in the last word is an entirely different one since the final fall is fully omitted. Is truncation then a case of elision of the final target tone? In a discussion on whether truncation is of phonological or rather phonetic nature, Grabe negates this interpretation for a number of reasons. The contours of <*Schiefer*> and <*Schief*> appear to prove the existence of a low target tone, which can be used as an argument against elision. If the final low target actually had to be deleted, the contour on <*Schief*> would have to be realized without it as well. A further reason is that the contour of <*Schiff*> resembles a possible falling contour of a longer utterance. Furthermore, the lack of perceptual difference between the three contours is another argument.

To summarize: this means that when the amount of voiced segments is reduced, truncation and compression are used by raising or lowering the rate of F0 change. Compression is indicated by a raise and truncation not by a steady state, but a lowering of the rate of change. This can be seen in Figure 3.10 for falling contours in German.

Concluding from this experiment, Grabe observes that the adjustment effects in German affect the entire final melody (pitch accent to boundary tone), although it is determined only by the pitch accent (rise: compression; fall: truncation). Therefore, a follow-up study was conducted to show that the choice of boundary tone has no influence on the adjustment strategy.

Experiment II

The second experiment focusses on the adjustment of H*+L in combination with the boundary tones H% (rise) and 0% (unspecified). This experiment offers no comparison of languages, but investigates compression in German intonation only. The founding hypothesis is that if compression occurs irrespective of the boundary tone specification, it is determined by the nuclear accent, rather than the nuclear tone.

The foundation of the experiment was, again, a corpus of read speech with a choice of materials that would evoke contours of the types L*+H H% and L*+H 0%. In order to obtain *coordinate structures* (see sect. 3.2.2), lists from a modified version of "Little Red Riding Hood" were used. Expected contours were L*+H 0% for lists and L*+H H% for questions.

Target words were:

<*Fisch*>, <*Fleisch*>, <*Fleischwurst*>

First, Grabe tried to confirm whether the lists and questions were actually constantly realized with the same boundary tone. After that, lists and questions were compared with respect to compression and, finally, acoustic measurements were made to identify H%-phrases and determine the rate of F0 change.

The result of that analysis was that all test words were realized with L*+H. The assumption, however, that lists would have the boundary tone 0% and questions H% could not be confirmed. Instead, that choice appears to be speaker-specific. Statistic measurements also did not result in any significant effects for H% and 0%. So, instead of investigating lists and questions in separate analyses, they were rearranged according to boundary tone specification. A summary of the results for duration, average slope and rate of F0 change can be seen in Figure 3.12.

The effects for duration are similar to those from the first experiment. The word containing the shortest voiced seg-

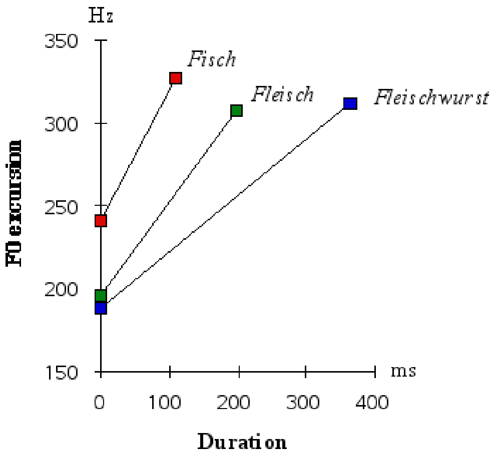

Fig. 3.12: Summarized results of the second experiment for the realizations of rising contours (x-axis: duration in ms; y-axis: F0 in Hz; rate of F0 change in Hz/ms) (Grabe 1998a, ch. 5.4.2.2).

ment (<*Fisch*>) is the shortest altogether and has the greatest slope. The longest word is, as expected, <*Fleischwurst*> and has the smallest slope. However, the level at which the contours end may suggest compression since it is almost the same for all three words. Therefore, it can be concluded that adjustment effects affect the nuclear accent and not the nuclear tone, i.e. H*+L and not H*+L H%/0%.

3.2 Method

3.2.1 Hypotheses

Although the present paper is based on prior research in the field of German intonation, particularly that stating truncation and compression as typical effects of it, it was not conducted using definite hypotheses but general assumptions about an underlying question: What happens to an intonation contour when only few voiced segments are available?

The hypotheses and results gained by several authors have already been mentioned in the introduction. However, as was shown in the introductory section, Hanssen et al. (2007) made the observation that speakers actually mix the two priorly assumed strategies. That idea as well as the possibility that the degree of interaction is largely speaker-dependent seems rather likely than that of two clearly separable and predictably occuring strategies. Can adjustment strategies really be reduced to truncation and compression or are there perhaps more than these two mechanisms that are used when only a short stretch of voicing is available? One important aspect is therefore, to find out how consistent inter-speaker realizations are and also whether intra-speaker inconsistencies are responsible for some degree of variation.

It is to be assumed, however, that falling contours do in fact behave differently than rising contours, given *declination* on the one, and *final lowering* on the other hand. *Declination* already decreases the final accent of a phrase (compared to a non-final accent), therefore an additional decrease in frequency to mark sentence type "statement" may be unnecessary (irrespective of voiced-unvoiced segments). The effect of *final lowering*, i.e. the lowering of the last pitch accent in a phrase, if it is a fall (Féry 1993, p. 56), may lead to the effect of a less steep fall being sufficient, and by that also a smaller rate of F0 change, to indicate a falling contour.

3.2.2 Materials

The present study was conducted with text material, compiled following Grabe's study. It consists of two different stories: a modified version of "Little Red Riding Hood" and a fictional dialogue about a lottery winner, triggering the following target words:

<*Mann*>	<*Nachbar*>	
<*Schiff*>	<*Schief*>	<*Schiefer*>
<*Fisch*>	<*Fleisch*>	<*Fleischwurst*>

The target words were embedded into the stories in phrase-final positions and occurred in statements, questions and lists. Since each word had to be read in every position once, a full set included three versions of each text. In addition, the speakers were asked to repeat the sets five times and to read in a normal speaking style to avoid a type of reading that is used when a fairy tale (like "Little Red Riding Hood") is read to a child, which may include the speaker paying special attention to his articulation and using an exaggerated intonation.

Although the texts were derived from Grabe (1998a,b), slight changes were conducted. Since the "Little Red Riding Hood" version initially used by Grabe served the purpose of also establishing an intonation inventory of the German language, it was simply too long for the questions raised in the present study and, therefore, reduced to those sentences containing the target words mentioned above. To these sentences, an introductory and a final sentence were added as well as another phrase: the original version contained statement and list only, but did not include questions, which were, therefore, added offering a full set of sentence types (statement list, question). The dialog experienced a change of a different nature since it already contained question and statement: it affected the socalled "coordinate structures" (Grabe 1998a), a term describing the event of similar contours occuring in two neigboring phrases. Grabe tries to include them by a content repetition as in *"Ist das nicht Herr Schiff? Unser neuer Nachbar?"* (engl. version: *"Isn't that Mr. Schift? Our new neighbor?"*). Theoretically, the contour realization on the mostly voiced word is underlyingly the same as that on the word with fewer voiced segments Grabe (1998a,b). Here, the text was modified by introducing a second counterpart for these *coordinate structures*, namely one that is monosyllabic. This was considered necessary because among the target words were also monosyllabic ones and for better comparison, a fully voiced single syllable opposition seemed called for: <*Mann*> (*man*).

In "Little Red Riding Hood" the *coordinate structures* are supposedly brought about by the phrase: *"Will sie wirklich*

Brombeermarmelade? Und Fisch?" (*"Does she really want blackberry jam? And fish?"*). However, a first inspection of the data already showed that the effect suggested by *coordinate structures* could not be found: The contours on the six-syllable words were completely different to those on bi- or monosyllabic ones. Thus, these words (all words containing *marmelade*) were not included in any further analysis.

3.2.3 Subjects

Ten subjects from northern Germany participated in the experiment, five female and five male. All were students at the Christian-Albrechts-University Kiel and between the age of 23 and 36.

3.2.4 Analyses

The recordings were made at the Institute of Phonetics and Digital Speech Processing (IPDS) in a highly sound-absorbent booth using a *Microtech Gefell M-940* large membrane condenser microphone and an *RME Fireface 800* soundcard at 24 bit amplitude resolution and 32 kHz sampling frequency.

Acoustic parameters analyzed, were duration and fundamental frequency (converted into semitones). Phone and word segmentation was carried out manually. The fundamental frequency was extracted by means of ESPS *get_f0* (Talkin 1995) in 10 ms steps and then stylized by polynomial first order regression or median filtering of the five first and last F0 values of a contour. All data were processed with *Matlab* followed by statistical analyses carried out using *SPSS 15*. Independent variables were: *F0 beginning*, *F0 end*, and *F0 delta*. The last variable describes the fundamental frequency's rate of change during a word and by that the contour's slope. Dependent variables were: *sentence type* ((ST) statement, question, list), *gender* (male, female), *speaker*, *word*, and *repetition*.

3.3 Results

3.3.1 Duration

Figure 3.13 shows clear differences in word length dependent vowel duration: <*Schiff*>, <*Fisch*>, and <*Mann*> are in all cases shorter than <*Schief*>, <*Fleisch*>, <*Schiefer*>, <*Fleischwurst*>, and <*Nachbar*>; <*Schief*> and <*Fleisch*> are always longer than <*Schiff*> and <*Fisch*> and at the same time shorter than <*Schiefer*>, <*Fleischwurst*>, and <*Nachbar*> (see Fig. 3.13).

As Fig. 3.13 shows, all words have greater length in questions than in statements. Phrased differently: Duration increases when the contour rises, but decreases when it falls. For the list realizations no clear observations of the kind could be made. In contrast to the other two sentence types, the longest word also had the longest duration, but in the

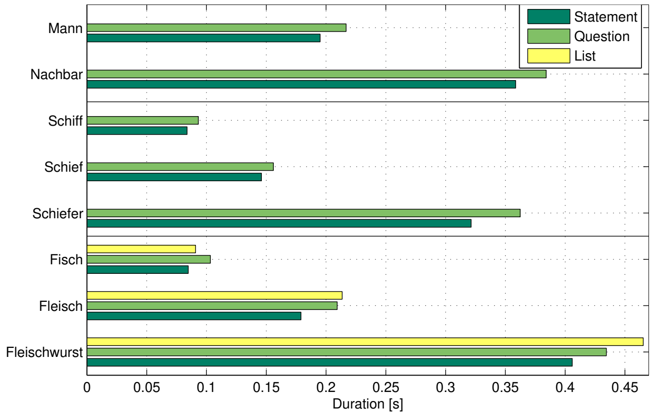

Fig. 3.13: Average duration values (in s) of target words: all speakers.

case of <*Fleisch*>, the difference is only minimal and entirely absent in the shortest word <*Fisch*>, with duration values between question and statement.

When regrouping the speakers according to gender, it can be seen that the duration of questions is generally greater than that of statements. Interestingly, both groups show the same inconsistencies when comparing the list variants to realizations of the other two sentence types. The female subjects realized the list form of the two longest words as also the longest sentence type variant. Only the durations of the shortest words lie between question and statement. The male speakers' realizations, on the other hand, mirror what has already been summarized for all speakers taken together.

To support Fig. 3.14 and 3.15, Tab. 3.1 offers a numerical summary of the durations of the sentence type relations for male and female subjects respectively. The number of events is the sum of the average duration values from all test subjects. Sentence types are marked by S for statement, Q for question and L for list and the duration relations by ">" for "longer than", "<" for "shorter than", and "=" for "equal to".

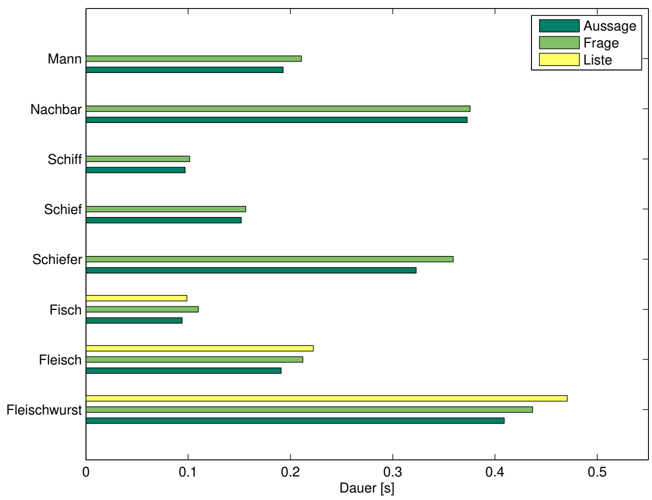

Fig. 3.14: Average duration values (in s) of target words: female speakers.

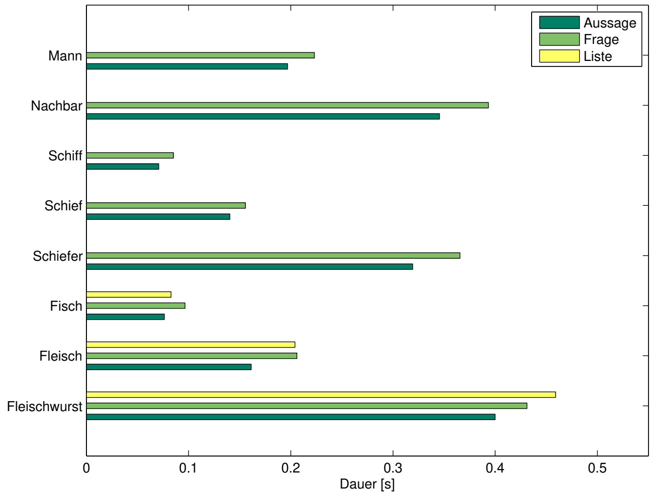

Fig. 3.15: Average duration values (in s) of target words: male speakers.

Due to the relatively small number of cases, absolute figures instead of percentages are given. The larger amount of statement-question-relations (40) results from the fact that each word occurs as a question as well as a statement (5 subjects * 8 words), while comparisons for list items could only be made for <*Fisch*>, <*Fleisch*> and <*Fleischwurst*> (therefore only 15 cases: 5 subjects * 3 words).

In general, male speakers produced very constant duration relations. Female speakers, on the other hand, produced relatively large inconsistencies, although favored realizations could nevertheless be identified. For male and female speakers, questions were in most cases longer than statements (70% and 90% respectively). In 25% of all cases, the female subjects produced longer statement than question variants. One speaker, among the male and female subjects each, produced exact same duration relations for questions and statements. In comparing list and statement variants it can be stated that duration realizations were independent of gender and that statement variants were generally longer.

Of course, conclusions about realizations of intonation categories cannot be drawn from duration values alone. Since the fundamental frequency is generally viewed as the acoustic correlate of intonation, the results of the F0 analysis will be introduced now.

Relation	S<Q	S>Q	S=Q	Q<L	Q>L	Q=L
Female	28	11	1	8	7	0
Male	36	3	1	8	6	1

Relation	S<L	S>L	S=L
Female	11	2	2
Male	13	2	1

Table 3.1: Numbers of duration relations of sentence types averaged over all female and male speakers (S: statement, Q: question, L: list).

3.3.2 F0 analysis

This section attempts to provide an answer to the question raised earlier: "How do phonetic realizations of phonological intonation categories in words with only little voiced material differ from realizations in words that are fully or mostly voiced?" Important parameters are *rate of change* and *slope*, i.e. the relation of beginning and final frequency of the F0 contour. First, the results from an analysis concerning the F0 delta values will be given, followed by results from the slope analysis.

3.3.3 Rate of change

The rate of change is specified by the delta value and expresses by how many semitones the contour changes within a word (or within the duration of the voiced segments). The average duration values of all speakers and all words are given in Fig. 3.16.

However, Tab. 3.2 presents all values of the rate of change in absolute numbers in order to allow for comparing the rates of change of the sentence types, irrespective of whether the change occurs into the negative or positive.

The figure and table, taken together, show that the three word groups (<*Mann*> and <*Nachbar*>; <*Schiff*>, <*Schief*> and <*Schiefer*>; <*Fisch*>, <*Fleisch*> and <*Fleischwurst*>) behave differently regarding the rate of change according to word length. While in the case of <*Mann*> and <*Nachbar*>, the rate of change increases for statements and decreases for questions (see Fig. 3.16), <*Schiff*>, <*Schief*> and <*Schiefer*> act differently: With increasing word length, the rate of F0 change increases for questions and statements. The results for <*Fisch*>, <*Fleisch*> and <*Fleischwurst*> however, display yet a different scheme with questions and statements, showing a similar pattern: The shortest word has the lowest rate of F0 change, while the highest rate can be found with the word <*Fleisch*>; in between lies the longest word of this group. For the statement variants, <*Fleisch*> is again dominant, since it has the lowest rate of F0 change. The highest rate can be observed for the shortest word <*Fisch*> and the longest word lies between the two.

For the parameter "F0 delta" an analysis of variance (univariate ANOVA) with the factors *sentence type* (*here:* state-

Word:	Mann	Nachbar	Schiff	Schief	Schiefer
S:	8	9	5	10	14
Q:	15	10	6	9	10

Word:	Fisch	Fleisch	Fleischwurst
S:	7	4	6
Q:	7	12	10
L:	1	5	4

Table 3.2: Rate of change/word depending on sentence type (S: statement, Q: question, L: list) in semitones.

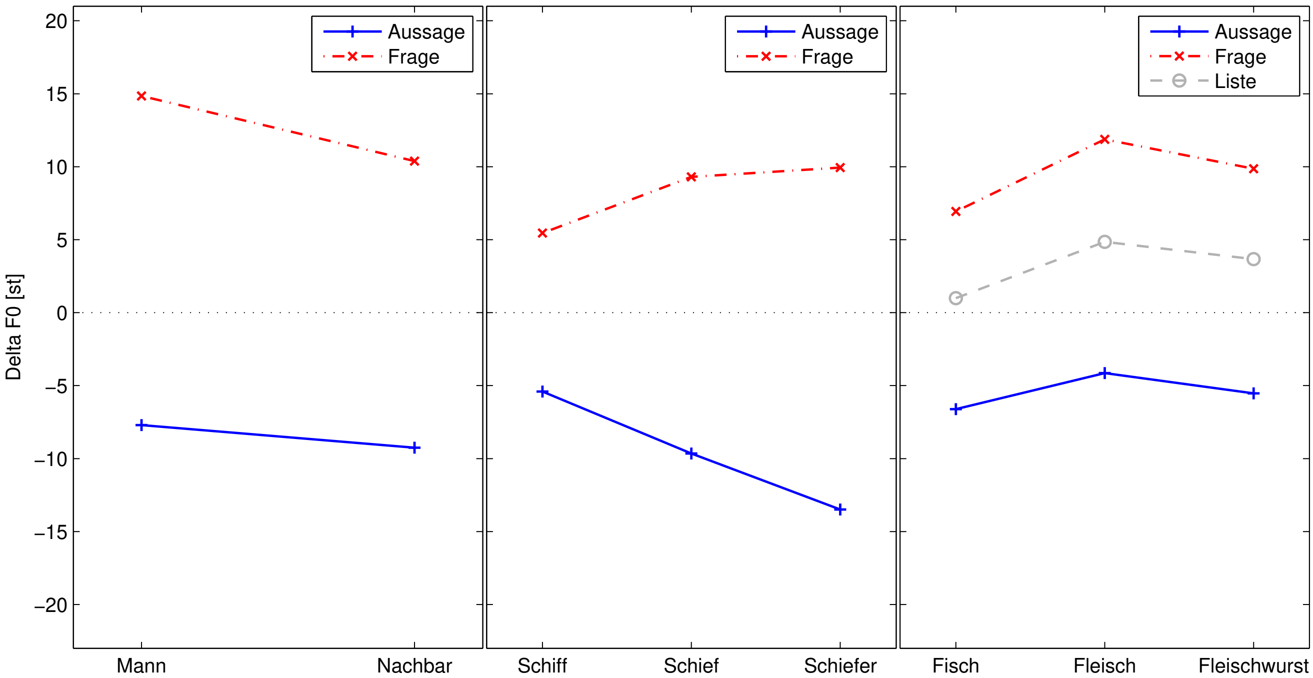

Fig. 3.16: Average F0 variation (in semitones) in all target words and for all sentence types.

ment and question), *word* (*here:* all) and *gender* was carried out. The results are given in Tab. 3.3 below.

Of the two main effects, *sentence type* has, by far, the greatest effect on F0 delta with an explained variance of 68.3%.

In order to include all sentence types in the statistics, another analysis of variance was carried out, using the same factors as before, but, additionally, containing all *sentence type* variants, i.e. statement, question, and list. The factor *word* is limited to <*Fisch*>, <*Fleisch*>, and <*Fleischwurst*> since only these words include the list variant. The results are presented in Tab. 3.4.

Here again, *sentence type* has the greatest effect on the rate of change accounting for 67.5% of the explained variance. Another 4.6% of the explained variance come about by the factor *word*. Since significance was found for the factor *sentence type*, a post-hoc test with Bonferroni correction was made. This test showed significant differences ($p<0.05$) in all cases, meaning that question, statement, and list differ from each other significantly.

Finally, the results for slope, i.e. the difference between initial and final frequency, will be presented, adding to the analyses previously introduced.

3.3.4 Slope of F0 contour

Beside the rate of F0 change, the slope, i.e. the rise or fall of the fundamental frequency, is an important component when studying intonation. Fig. 3.17 shows a combination of average duration values and rates of change, measured for the various word lengths. The initial value for every word in every position is set to zero since the figure does not display absolute frequency values but a stylized course of the F0 contours. This figure shows what happens to F0 contours in different sentence types and how these interact with each other.

F0 delta	df	*F*	p	explained variance
ST	1	4571.900	<0.001	68.34%
Word	7	682.444	<0.001	3.82%
Gender	1	41.724	0.135	0.03%
ST*Word	7	894.917	<0.001	5.01%
ST*Gender	1	1093.700	<0.001	0.88%
Word*Gender	7	18.989	0.418	0.11%
ST*Word*Gender	7	221.848	<0.001	1.24%

Table 3.3: Results of the analysis of variance for F0 delta with the factors *sentence type* (ST *here*: statement, question), *word* (*here*: all), and *gender*.

F0 delta	df	*F*	p	explained variance
ST	2	551.178	<0.001	67.54%
Word	2	37.449	<0.001	4.59%
Gender	1	5.044	0.025	0.31%
ST*Word	4	2.475	0.044	0.61%
ST*Gender	2	15.158	<0.001	1.86%
Word*Gender	2	2.306	0.101	0.28%
ST*Word*Gender	4	4.236	0.002	1.04%

Table 3.4: Results of the analysis of variance for F0 delta with the factors *sentence type* (*here*: statement, question, list), *word* (*here*: <*Fisch*>, <*Fleisch*>, <*Fleischwurst*>) and *gender*.

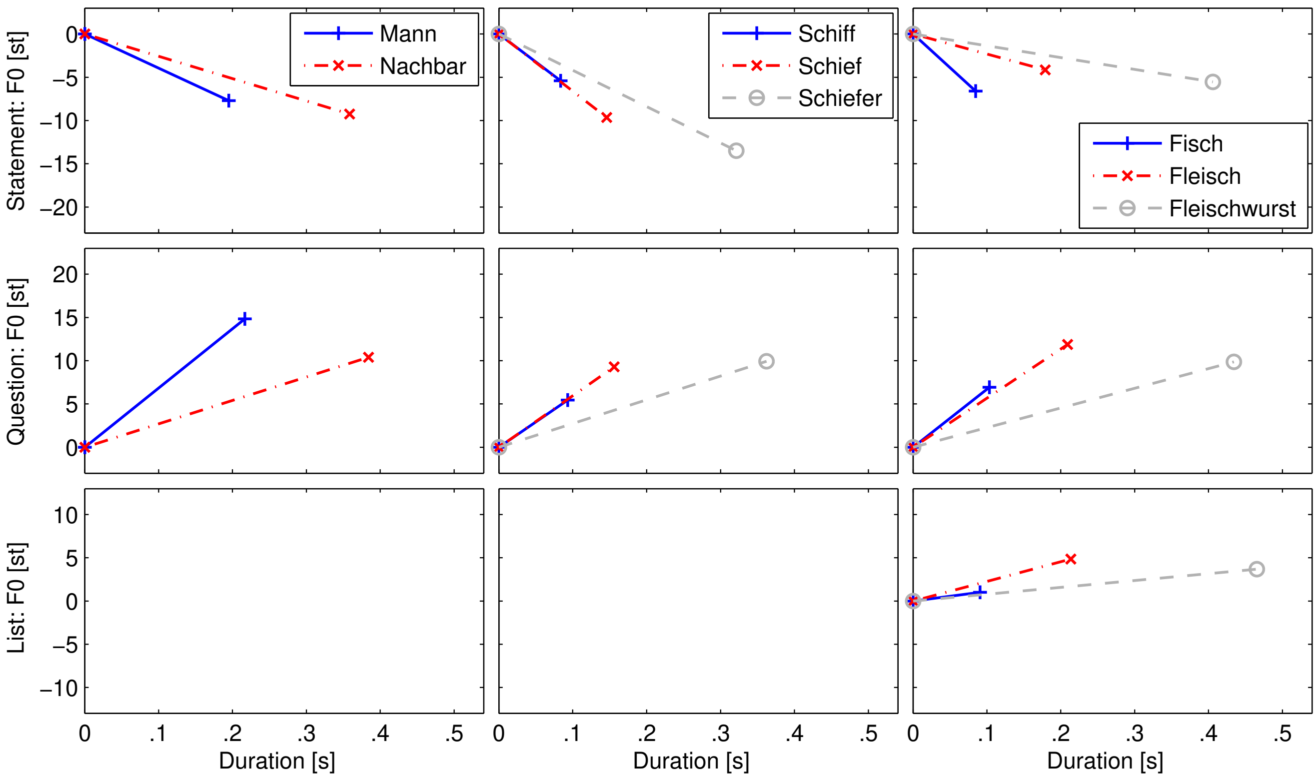

Fig. 3.17: Average slope (st/s) in the target words for all test subjects.

The results shown in Fig. 3.17 are, as already mentioned, a summary of duration and delta values. Since these two are presented in the same figure, the results can be used to add to and confirm previous observations. The delta value shows the rate of change of semitones per word. Since the words are not equally long, a direct comparison of absolute numbers is not ideal. However, Fig. 3.17 offers a good basis for such a comparison: Words with a segmentally small amount of voiced material are also the shortest; long words display a smaller slope in questions and statements. The similar behavior, when realizing the phonological categories H*+L (statement) and L*+H (question, list) in short words also makes this apparent. <*Fisch*> and <*Mann*> always have the greatest slope in questions and statements. In lists, the slope of <*Fisch*> equals that of <*Fleischwurst*>. <*Schiff*> does not stick out because of its great slope in questions and statements, but because of sharing the exact same slope as <*Fleisch*>.

The contour is more even for list realizations of all words than the slope in questions. Differences between statements and questions, concerning the slope of the contours, are minimal and inconsistent. <*Mann*>, <*Nachbar*>, <*Fisch*>, <*Fleisch*>, and <*Fleischwurst*> have more even contours in statements than in questions. The slope of <*Fisch*>, <*Fleisch*>, and <*Fleischwurst*> is similar in statements and lists. In the case of <*Schiff*>, <*Schief*>, and <*Schiefer*>, the relation of statements and lists is almost exactly the other way around. Statement contours have a greater slope than question contours.

In the following, statistic analyses that give an overview of the influence of several factors on the shape of a contours' slope will be presented. In order to carry out the statistic analyses for the slope, "F0 beginning" and "F0 end" were chosen as dependent variables. A summary of the most important results will follow in a short description. An inter-

F0 beginning	df	*F*	p	explained variance
ST	1	2230.124	<0.001	18.66%
Word	7	226.587	<0.001	13.27%
Subj	9	625.922	<0.001	47.12%
ST*Word	7	49.762	<0.001	2.91%
ST*Subj	9	42.045	<0.001	3.17%
Word*Subj	63	3.669	<0.001	1.93%
ST*Word*Subj	63	4.747	<0.001	2.50%

F0 end	df	*F*	p	explained variance
ST	1	4511.281	<0.001	46.66%
Word	7	49.999	<0.001	3.62%
Subj	9	395.301	<0.001	36.80%
ST*Word	7	64.141	<0.001	4.64%
ST*Subj	9	26.147	<0.001	2.43%
Word*Subj	63	4.004	<0.001	2.61%
ST*Word*Subj	63	4.982	<0.001	3.25%

Table 3.5: Results from the analysis of variance for F0 beginning (top) and F0 end (bottom) with the factors *sentence type* (ST, *here*: statement, question), *word* (*here*: all) and *test subject*.

F0 beginning	df	*F*	p	explained variance
ST	2	72.754	<0.001	5.13%
Word	2	110.658	<0.001	7.81%
Subj	9	211.974	<0.001	67.30%
ST*Word	4	11.692	<0.001	1.65%
ST*Subj	18	6.418	<0.001	4.08%
Word*Subj	18	1.438	0.112	0.91%
ST*Word*Subj	36	1.559	0.025	1.98%

F0 end	df	*F*	p	explained variance
ST	2	882.272	<0.001	37.97%
Word	2	3.807	0.023	0.16%
subj	9	244.311	<0.001	47.32%
ST*Word	4	7.648	<0.001	0.66%
ST*Subj	18	12.007	<0.001	4.65%
Word*Subj	18	2.383	0.001	0.92%
ST*Word*Subj	36	2.021	0.001	1.57%

Table 3.6: Results from the analysis of variance for F0 beginning (top) and F0 end (bottom) with the factors *sentence type* (ST, *here*: statement, question, list), *word* (*here*: <*Fisch*>, <*Fleisch*>, <*Fleischwurst*>) and *test subject*.

pretation of those results will be presented in a later section.

For F0 beginning, all effects and interactions of the factors are significant. The biggest effect with an explained variance of 47.1% is *test subject* (Subj). Only 18.7% of the variance can be explained by *sentence type* and another 13.3% by *word*. For F0 end, also all effects and interactions are significant, but the 46.7% explained variance of *sentence type* exceed the 36.8% of *test subject* (see Tab. 3.5).

The effects for F0 beginning and F0 end of the three sentence types are quite similar to those mentioned earlier. F0 beginning has significant effects and interactions apart from one: *word* and *test subject*. For F0 end, every factor is significant. The fact that 67.3% of the variation of F0 beginning can be accounted for by *test subject* is quite noteworthy. However, *sentence type* explains 38.0% and *test subject* another 47.3% of the variance of F0 end. Also significant and with an explained variance of 4.7% is the interaction of *sentence type* and *word* (see Tab. 3.6).

3.4 Discussion

The summary of results will be given in the same order as has been used previously: duration, rate of change and slope.

3.4.1 Duration

It was shown that words were adjusted depending on sentence type, being the longest for questions (see Fig. 3.13). This was particularly obvious when contrasting the duration values of statements and questions (averaged over all speakers) since questions were always the longest. Comparing questions and lists did not lead to such a clear result because the two do not share such a systematic relation. A reoccurring pattern could, however, be observed in the comparison of list and statement: Statement variants are always shorter than list variants. This could not be accounted for by *gender* but appears to be much rather a speaker-specific choice.

What does this mean for the underlying question? Since utterances of the sentence type question take the longest, one could also say: In the case of rising contours, the underlying segmental material is lengthened, while in the case of falling contours, a shorter duration is sufficient. However, duration in itself is not the only parameter that determines sentence type.

3.4.2 Rate of F0 change

An even course of the rate of F0 change could only be observed for the word group <*Schiff*>, <*Schief*>, <*Schiefer*>. Within this group, the rate of change increased proportionally to the growing amount of voiced segments in falling and rising contours. In case of the sentence types statement and question, an analysis of variance of the dependent variable F0 delta showed a main effect for *sentence type* with an explained variance of 68.3%. Another main effect of *sentence type* occurred at a comparison of all sentence types (statement, question and list) with an explained variance of 67.5%. A post-hoc test showed also that the three sentence types differed significantly.

For the present paper, these results confirm that statement and question contours represent two different phonological categories: H*+L, i.e. falling, and L*+H, i.e. rising. They also allow to draw conclusions about the shape of question and statement F0 contours since they differ significantly with respect to F0 delta (rate of change).

3.4.3 Slope

The slope combines aspects of duration and rate of change and, in that, offers a good summary of the results altogether (see Fig. 3.17). An analysis showed that the shortest words were generally also those with the greatest slope, not depending on whether it was a question, a statement or a list item. The segmentally longest words had, in most cases, the smallest slope and the words of medium segmental length (*here:* <*Fleisch*>, <*Schief*>) displayed no common adjustment strategy. While on average of all speakers, <*Schief*> had the same slope as <*Schiff*>, the realizations of <*Fleisch*> vary greatly.

In order to compare these results to those of Grabe (1998a,b), see Fig. 3.18.

When looking back to the contours of the slopes measured from the data of the present study, a different figure must result.

Grabe (1998a,b) as well as Hanssen et al. (2007) use slope, i.e. the increase or decrease of the rate of F0 change, as a correlate of truncation and compression. If the rate of change increases upon a shortening of the amount of voiced

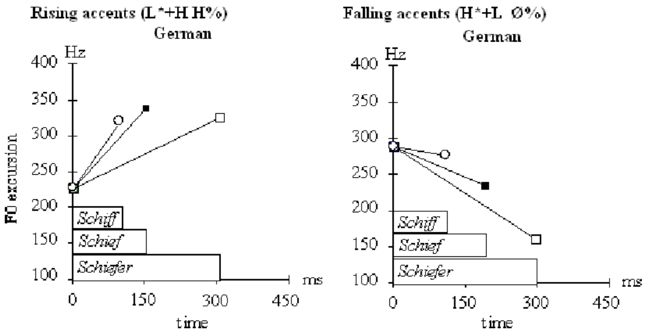

Fig. 3.18: Results of F0 slope for rising (left) and falling contours (right) (Grabe 1998a, ch. 5.3.2).

material, compression is assumed. Truncation is only assumed, if the rate either remains constant or decreases.

Before one tries to derive general results from the figure of slopes (see Fig. 3.17), the results from the statistical analyses should be considered. The ANOVA results (Tab. 3.5) for two and three sentence types of the dependent variables F0 beginning and F0 end lead to similar variances in both cases: For the two sentence types and F0 beginning, *test subject* made up 47.1% of the explained variance and in the case of F0 end 46.7% of the explained variance were *sentence type* and another 36.8% *test subject*. For the three sentence types, a main effect for *test subject* showed for F0 beginning with an explained variance of 67.3% and for F0 end, *test subject* (47.3%) and *sentence type* (38.0%).

With regards to Grabe's results, this indicates that speaker-specific realization preferences can not be ruled out and must obviously be taken into account as well.

The data of the test subjects reveal that the choice of realization strategy is strongly speaker-dependent and not a characteristic of an entire language system (*here:* German). While the slopes of the words show intra-speaker consistency, another speaker may choose an entirely different strategy. The difference between initial and final frequency of an F0 contour varies from speaker to speaker and also accross different word groups.

3.5 Conclusion

Of course, in order to allow for generalizations, further research is needed: studies with more test subjects, more word groups and especially with data coming not exclusively from read speech but also from spontaneous speech.

This was already the goal of a second experiment of which there was no mention so far. Initially, it was intended to shed light on phonetic realization strategies in spontaneous speech, using data gained from the *Video Task Scenario: Lindenstraße* of the *Kiel Corpus of Spontaneous Speech*. This corpus was designed at the IPDS Kiel, containing six dialogs among four female and two male couples and amounting to a total duration of 80 minutes and 13.000 words (Peters 2005). The reason for no detailled description of that second experiment was that only very little data matching the criteria required for further analyses, was available in the corpus (8 phrase-final statements and 2 phrase-final questions). So no conclusive observations could be made from that corpus.

However, a more thourough analysis of spontaneous speech seems relevant, for only these cases can show the use of adjustment strategies in naturally spoken language.

Of course, the results from the present paper already indicate that the existence of two strategies used in German intonation is not as certain as it seemed. The data showed no reoccuring strategies used by all speakers at all. Rather, each speaker appeared to have his own strategy, which he used consistently, wherefore, a study with more test subjects would be required and perhaps even comparisons of different dialect areas, since relevance of that parameter was already indicated by Bannert & Bredvad-Jensen (1975). Another important factor seems to be the word group in itself, as it was shown that each word group behaved differently.

Bibliography

Adriaens, L. M. H. (1991). *Ein Modell deutscher Intonation. Eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text.* Ph. D. thesis, Eindhoven University.

Atterer, M. & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *J. of Phonetics 32*, 177–197.

Bannert, R. & Bredvad-Jensen, A.-C. (1975). Temporal organization of Swedish tonal accents: The effect of vowel duration. Working Papers 10, Phonetics Laboratory, Department of General Linguistics, Lund University, Lund, 1–36.

Bierwisch, M. (1966). Untersuchungen über Akzent und Intonation im Deutschen. In: Bierwisch, M. (Ed.), *Studia Grammatica VII*. Berlin: Akademie-Verlag, 99–201.

Dilley, L., Ladd, D. R. & Schepman, A. (2005). Alignment of L and H in bitonal pitch accents: testing two hypotheses. *J. of Phonetics 33*, 115–119.

Erikson, Y. & Alstermark, M. (1972). Fundamental frequency correlates of the grave word accent in Swedish: the effect of vowel duration. Speech Transmission Lab., Quarterly Progress and Status Report 2–3, KTH, Stockholm, 53–60.

Féry, C. (1993). *German Intonational Patterns*. Tübingen: Niemeyer.

Fujisaki, H. (2004). Information, prosody, and modeling with emphasis on tonal features of speech. In: *Proc. of the 2nd Int. Conf. on Speech Prosody*. Nara; Japan, 1–10.

Gartenberg, R. & Panzlaff-Reuter, C. (1991). Production and perception of F0 peak patterns in German. Arbeitsberichte (AIPUK) 25, IPDS, Christian-Albrechts-University, Kiel, 29–113.

Grabe, E. (1998a). *Comparative intonational phonology: English and German*. Ph. D. thesis, Nijmegen University, Wageningen.

Grabe, E. (1998b). Pitch accent realization in English and German. *J. of Phonetics 26*, 129–143.

Hanssen, J., Peters, J. & Gussenhoven, C. (2007). Phrase-final pitch accommodation effects in Dutch. In: *Proc. ICPhS*, vol. 2. Saarbrücken, 1077–1080.

Isačenko, A. V. & Schädlich, H.-J. (1966). Untersuchungen über die deutsche Satzintonation. In: Bierwisch, M. (Ed.), *Studia Grammatica VII*. Berlin: Akademie-Verlag, 7–67.

Isačenko, A. V. & Schädlich, H.-J. (1970). *A model of standard German intonation*. The Hague: Mouton.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Niebuhr, O. & Ambrazaitis, G. (2006). Alignment of medial and late peaks in German spontaneous speech. In: *Proc. of the 3rd Int. Conf. on Speech Prosody*. Dresden, 161–164.

Niebuhr, O. & Pfitzinger, H. R. (2010). On pitch-accent identification. The role of syllable duration and intensity. In: *Proc. of the 5th Int. Conf. on Speech Prosody*. Chicago.

Ohl, C. K. & Pfitzinger, H. R. (2009). Compression and truncation revisited. In: *Proc. Interspeech 2009*. Brighton, 2451–2454.

Peters, B. (2005). The database — the Kiel Corpus of Spontaneous Speech. Arbeitsberichte (AIPUK) 35a, IPDS, Christian-Albrechts-University, Kiel, 1–6.

Peters, B. & Pfitzinger, H. R. (2008). Duration and F0 interval of utterance-final intonation contours in the perception of German sentence modality. In: *Interspeech*. Brisbane, 65–68.

Pfitzinger, H. R. (2002). Intrinsic phone durations are speaker-specific. In: *Proc. of 7th ICSLP*, vol. 2. Denver, 1113–1116.

Pfitzinger, H. R., Mixdorff, H. & Peters, B. (2009). Correspondences between KIM-based symbolic prosodic labels and parameters of the Fujisaki model. In: *Proc. Nordic Prosody X*. Helsinki, 261–272.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Ph. D. thesis, Indiana. MIT Dissertation.

Rapp, S. (1998). Automatisierte Erstellung von Korpora für die Prosodieforschung. Working papers 4(1), Institut für Maschinelle Sprachverarbeitung, Stuttgart.

Talkin, D. (1995). A robust algorithm for pitch tracking (RAPT). In: Kleijn, W. B. & Paliwal, K. K. (Eds.), *Speech coding and synthesis*. New York: Elsevier, Ch. 14, 495–518.

Uhmann, S. (1997). *Fokusphonologie: eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.

von Essen, O. (1964). *Grundzüge der hochdeutschen Satzintonation*. Ratlingen, Düsseldorf: Henn.

Wunderlich, D. (1988). Der Ton macht die Melodie — Zur Phonologie der Intonation des Deutschen. In: Altmann, H. (Ed.), *Intonationsforschungen*. Tübingen: Niemeyer, 1–40.