Zusammenfassung

Für eine Vielzahl von Sprachen wurde gezeigt, dass eine dynamische f0-Kontur zu einer Verlängerung der wahrgenommenen Vokaldauer beiträgt und Hörer von diesem Parameter bei der Disambiguierung uneindeutiger Fälle Gebrauch machen. In der vorliegenden Studie wird der perzeptuelle Effekt am Beispiel trilingualer Saterfriesen untersucht, da die Ergebnisse früherer Produktionsstudien nahelegen, dass f0-Dynamik zusätzlich kontrastverstärkend für die Unterscheidung phonemischer Vokaloppositionen im Saterfriesischen wirken kann. Die Analyse der trilingualen Saterfriesen erfolgt kontrastiv zu monolingualen Sprechern des nördlichen Standarddeutschen, um eine mögliche sprachspezifische Ausprägung des perzeptuellen Effekts zu berücksichtigen. Es werden resynthetisierte natürliche Stimuli verwendet. Dabei handelt es sich um monosyllabische Kunstwörter der Form /hVt/ mit je drei Dauerstufen und zwei Konturverläufen (flache f0-Kontur vs. dynamische f0-Kontur) pro untersuchter Vokalkategorie. Für beide Sprechergruppen lässt sich ein Einfluss der f0-Kontur auf die wahrgenommene Vokaldauer beobachten. Stimuli wurden als länger wahrgenommen, wenn diese mit einer dynamischen anstelle einer flachen f0-Kontur realisiert wurden. Bei den trilingualen Sprechern ist dieser Effekt jedoch auf die halblangen und langen Stimuli geschränkt. Damit zeigt sich, dass eine dynamische f0-Kontur generell zu einer Verlängerung der wahrgenommenen Vokaldauer beiträgt, unabhängig vom sprachlichen Hintergrund der Sprecher. Durch die Beschränkung des Effekts auf die halblangen und langen Stimuli bei den trilingualen Sprechern deuten die Ergebnisse zugleich auf eine sprachspezifische Implementierung des perzeptuellen Effekts hin, welcher kontrastverstärkend zur Differenzierung von Kurz- und Langvokaloppositionen beitragen kann.

Wilbert Heeringa[a]
Fryske Akademy, PO Box 54, 8900 AB Ljouwert, The Netherlands, wheeringa@fryske-akademy.nl

Heike E. Schoormann
Institute of German Studies, University of Oldenburg, Ammerländer Heerstrasße 114-118, 26129 Oldenburg, Germany, heike.schoormann@uol.de

Jörg Peters
Institute of German Studies, University of Oldenburg, Ammerländer Heerstrasße 114-118, 26129 Oldenburg, Germany, joerg.peters@uol.de

[a] Corresponding author

Wilbert Heeringa, Heike Schoormann, and Jörg Peters

Monolingual and trilingual perception of duration in Saterland Frisian vowels

1. Introduction

Listeners with a different language background rely on vowel duration in distinguishing between short and long vowels, suggesting that vowel duration is a salient perceptual cue (BOHN 1995, cf. LEHNERT-LEHOUILLER 2010). In addition to vowel duration, listeners make use of secondary cues, such as f0, to disambiguate vowel quantity (cf. KINOSHITA, BEHNE & ARAI 2002, LEHNERT-LEHOUILLIER 2010). Cross-linguistically, dynamic f0, as can be found in contours with a falling, rising, rising-falling or falling-rising pattern, was shown to increase perceived vowel duration when compared to level f0 (cf. LEHISTE 1976, PISONI 1976, WANG, LEHISTE, CHUANG & DARNOVSKY 1976, YU 2010 on American English; CUMMING 2011 on Swiss German, Swiss French, and French). LIPPUS, PAJUSALU & ALLIK (2011) studied the role of pitch in the Estonian three-way quantity system and showed that it is possible to generate the perception of an overlong word from a long word by changing its pitch contour. LEHNERT-LEHOUILLIER (2010) studied languages with a phonemic vowel length contrast (German and Spanish) and languages with the contrastive use of f0 (lexical tone in Thai and pitch accent in Japanese) and found an effect of language background on the perceptual lengthening of a dynamic f0 contour. In her cross-linguistic study, only the Japanese listeners rated stimuli with a falling f0 as longer than stimuli with a level f0 contour. Earlier, VAN DOMMELEN (1991, 1993) had found a perceptual lengthening effect of dynamic f0 in German but only in isolated monosyllabic words.

Despite contradicting evidence, the previous studies have shown that dynamic pitch may have a lengthening effect but that this effect is likely to be language-specific and context-dependent. The observed lengthening effect seems to be independent of the presence or absence of a phonemic length contrast, although the most consistent findings were found for American English, a language without a phonemic vowel length contrast (cf. Lehnert-LeHouiller 2010). In addition, the investigated languages varied with regard to the number of distinguished vowel qualities, ranging from five for Spanish and Japanese to 15 for German and 17 for Swedish (cf. ŠIMKO, AALTO, LIPPUS, WŁODARCZAK & VAINIO 2015 on Estonian, Finnish, Mandarin, and Swedish), and with regard to whether differences in vowel length are linked to differences in vowel quality, as in Standard German, or not (cf. WIESE 2000, 21).

The present paper studies the perceptual lengthening effect of dynamic f0 in Saterland Frisian, spoken in the municipality of Saterland in northwest Germany. Saterland Frisian is the last remaining variety of East Frisian and one of the most endangered minority languages in Europe. Latest estimates of the number of native speakers range between 2,250 and 2,500 (STELLMACHER 1998, FORT 2015, XIII). Although the dialectal differences between the three local dialects spoken in Ramsloh, Strücklingen, and Scharrel/Sedelsberg are small (FORT 2015, 817, SCHOORMANN, HEERINGA & PETERS 2017a), native speakers are highly aware of regional differences in the pronunciation of Saterland Frisian. Of the three dialects, Ramsloh Saterland Frisian is reported as the most conservative and has maintained Old East Frisian features not found in the present-day dialects of Strücklingen and Scharrel (FORT 2015: XIV, 817). Due to extensive language contact with Low German and Northern High German, native speakers of Saterland Frisian are either bilingual with Saterland Frisian and the local variety of Northern High German, henceforth Saterland High German, or trilingual with Saterland Frisian, Saterland High German and Saterland Low German. The Low German variety spoken in the Saterland is a mixture of Münsterland and Emsland Low German (FORT 2004).

According to SJÖLIN (1969, 67), KRAMER (1982), and FORT (2015, XV), Saterland Frisian has a complete set of close short tense vowels: /i y u/. Together with the short lax vowels /ɪ ʏ ʊ/ and the long tense vowels /iː yː uː/ they constitute series of phonemes that differ by length and/or tenseness. In addition to the three-way distinction of close vowels, Saterland Frisian and Low German have short (/ɛ œ ɔ/) and long (/ɛː œː ɔː/) open-mid lax vowels, which stand in opposition to long tense vowels (/eː øː oː/). Figure 1 compares the three vowel systems of Saterland Frisian, Low German, and High German.

Saterland Frisian	Low German	High German
iː/i yː/y uː/u	iː yː uː	iː yː uː
ɪ ʏ ʊ	ɪ ʏ ʊ	ɪ ʏ ʊ
eː øː oː	eː øː oː	eː øː oː
ɛː œː ɔː	ɛː œː ɔː	ɛː/ œ ɔ
ɛ œ ɔ	ɛ œ ɔ	ɛ
a aː	a aː	a aː

Figure 1. The inventory of monophthongs of Saterland Frisian (left), Low German (middle), and High German (right).

SIEBS' (1966 [1889]) distinction between *Stoßton* ('pushing tone') and *Schleifton* ('dragging tone') in Saterland Frisian suggests that f0 might have played a role in the differentiation of the close vowels in the past. However, TRÖSTER-MUTZ (1997, 2002) did not find any evidence for a tone accent distinction in present-day Saterland Frisian (see also PETERS 2008). HEERINGA, PETERS & SCHOORMANN (2014) examined the cues that distinguish the close front unrounded and the back rounded series of short lax and short and long tense vowels in minimal triplets by eliciting 'normal speech' and 'clear speech' in a reading task. In the 'clear speech' condition, which was designed as a listener-directed task for maximum discrimination, a triplet word was always presented together with the two other triplet words. In this condition, f0 dynamics and centralization in the F1-F2 space were used as additional means to make short tense vowels more distinct from long tense vowels. These results suggest that both length and tenseness are used as distinctive features, while f0 dynamics and centralization in the F1-F2 space were optionally used to enhance the contrast between short and long tense vowels. Although SCHOORMANN, HEEERINGA & PETERS (2017a) found that the close short and long tense vowels have merged in present-day Saterland Frisian, the results of HEERINGA, PETERS & SCHOORMANN (2014) suggest that some older Saterland Frisian speakers are nonetheless still sensitive to f0 cues in the perception of vowel duration.

The objective of the present study is to examine the interaction between f0 cues and the listeners' language background on perceived vowel duration. To this end, the duration ratings of the trilingual speakers from the Saterland are compared to the ratings of monolingual speakers of Northern High German from outside the Saterland. In stressed syllables of Northern High German, vowel quality and tenseness correlate with vowel duration, i.e. tense vowels are long, and lax vowels are short. The only two exceptions are /ɛː/ and /ɛ/, and /aː/ and /a/, which show little or no spectral differences and primarily differ in acoustic duration (see Figure 1). In Northern High German /ɛː/ is largely restricted to spelling pronunciation and careful speech. In more informal speech, /ɛː/ tends to be merged with /eː/ (cf. BOHN & FLEGE, 1992, JØRGENSEN, 1969, KOHLER, 1995, 172-173, PÄTZOLD & SIMPSON 1997, STEINLEN, 2005, 79). On the other hand, Saterland Frisian and Low German have a series of long lax vowels, which differ from short lax vowels and long tense vowels in both duration and spectral features, i.e. /ɛː ɛ eː/, /œː œ øː/, and /ɔː ɔ oː/.

For the comparison with monolingual speakers from outside the Saterland we recruited speakers from Hanover. The rationale behind choosing speakers from Hanover as opposed to Northern High German monolinguals from the Saterland is that speakers from the Saterland and the immediate surrounding area, who consider themselves monolingual, are hardly ever truly monolingual. Instead, they often have at least passive knowledge of Low German, which might influence their perception. This is also true for the speakers of neighboring cities, such as Leer, Oldenburg, and Bremen. We chose participants from Hanover because they may be regarded as representative of speakers of northwestern Standard German who have had little contact with Low German. Local peculiarities of the Hanover variety have been on the retreat since the turn of the 20th century (cf. ELMENTALER 2012). Today's urban vernacular of Hanover hardly shows any regional characteristics and may be considered one of the local varieties of Northern High German which come close to the codified standard. In addition, Hanover High German is commonly considered most typical of Northern High German as used in the North German media.

Because vowel duration is the most salient and universal cue, we expect that our participants make equal use of this cue in distinguishing between long and short vowels. On the other hand, we expect that the language background may have an effect on the use of secondary cues such as f0 dynamics, which are less salient and less commonly used. In particular, we will examine the following questions:

(1) Are vowel stimuli of equal acoustic duration perceived as relatively shorter or longer when realized with a flat or a rising-falling f0 contour? And do listeners from the Saterland and Hannover show the same sensitivity to the f0 cue?

(2) Is there an interaction of vowel duration and f0 contour or does a dynamic f0 affect short, half-long, and long vowels equally? And do Saterland and Hannover listeners show the same interaction?

(3) Is there an interaction between f0 contour and vowel quality or is the lengthening effect of a dynamic f0 contour the same for all vowels? And do both groups of listeners show the same interaction?

2. Method

2.1 Subjects

We recruited 22 trilingual subjects from Scharrel, 11 male and 11 female speakers, aged between 51 and 75 years. All speakers were born and raised in Scharrel and had lived in Scharrel for all or the majority of their lives. They were trilingual with the local dialect of Saterland Frisian, Saterland Low German, and the local variety of Northern High German. All speakers acquired Saterland Frisian at home and reported to use it as their primary language in communication with other speakers of Saterland Frisian. Next to Saterland Frisian, the local variant of Low German was generally acquired in early childhood at home, in the neighborhood, and from relatives who did not speak Saterland Frisian. Most speakers learned High German when entering school. In addition, we recruited 11 male and 10 female monolingual speakers of Northern High German from the city of Hanover and its surrounding area. Hanover is the state capital of Lower Saxony and situated about 170 kilometers southeast from

Scharrel. Most of the speakers were raised in Hanover and had lived there for all or the majority of their lives. All monolingual subjects were aged between 50 and 74 years.

2.2 Eliciting representative acoustic stimuli

We examined the following ten Saterland Frisian vowels in the perception test: /iː/, /ɪ/, /uː/, /ʊ/, /eː/, /ɛː/, /ɛ/, /oː/, /ɔː/, /ɔ/. Since /iː i/ and /uː u/ were found to be merged in our subjects in a previous experiment (SCHOORMANN, HEERINGA & PETERS 2017a) the merged categories were pooled. Each vowel was recorded in a /hVt/ frame by a native speaker of the Scharrel dialect. In order to obtain representative vowel stimuli, we defined a model speaker based on the vowel productions of the 11 male speakers who had earlier participated in production tests by HEERINGA, SCHOORMANN & PETERS (2015) and SCHOORMANN, HEEERINGA & PETERS (2017a). To this end, we determined the group median F1 and F2 values for each vowel from the realizations of all speakers and selected the speaker with the smallest distance to the median F1 and F2 values averaged over all vowels.

/hVt/ words were cued by reading aloud rhyming Saterland Frisian words immediately preceding the production of the /hVt/ target word (cf. BOHN 2004). We preferred to elicit target words via rhymes over a usual reading task because Saterland Frisian orthography was unknown to our speaker and the written form may have had a direct influence on the reading task. As trigger words we used monosyllables ending with [t]. For example, in order to obtain Saterland Frisian /eː/, first the Saterland Frisian word *leet* 'late' was shown, together with its High German translation. The model speaker then read out /leːt/ aloud and subsequently built the rhyming target word /heːt/ from the *H_t* frame provided on the screen. Where no monosyllabic triggers ending on [t] were available, an intermediate form was shown between the trigger and the target word. For example, in order to obtain the Saterland Frisian /œ/ from the trigger *löskje* 'extinguish', the intermediate form *lött* was added to elicit the rhyming target word /hœt/. Recordings were preceded by practice words so that informants became familiarized with the task. Each sequence of the trigger and target word was presented six[b] times per vowel in controlled randomized order.

Mean F1 and F2 values were calculated on the basis of the multiple repetitions by the model speaker in order to obtain the most representative instance per vowel category.[c] Similar to the detection of the model speaker, the most representative instance per vowel category was the repetition with the smallest average distance to the mean F1 and F2 value. Striking is the strong qualitive distinction between /ɛː/ and /ɛ/, and the low F1 values of /ɪ/ and /ʊ/.

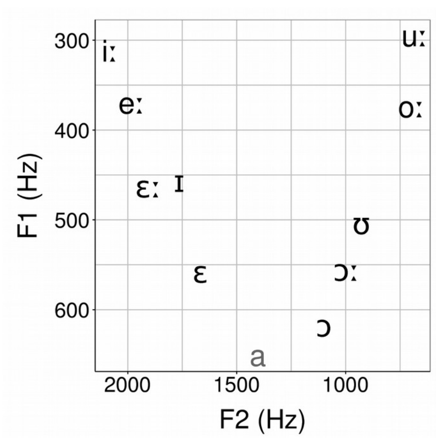

Figure 2. Most representative instances of the ten vowel categories in the F1-F2 space, pronounced by a representative speaker of Saterland Frisian in Scharrel. The /a/ is added for visualization purposes only.

2.3 Generating stimuli for the perception experiment

The most representative vowel utterances determined in 2.2 were used to create stimuli of three different durations – short, half-long, and long – and with two different f0 contours – a flat contour and a rising-falling (dynamic) contour – for each vowel category. Half-long stimuli were included since short tense vowels are often described as half-long vowels and also tend to be realized with slightly longer acoustic durations than the short lax vowels (cf. KRAMER 1982, 5, FORT 2015, XV, HEERINGA,

b Because of a recording error only five repetitions were obtained for /ɔ/.

c We recorded 12 repetitions for /uː/, 13 for /iː/, and seven for /ʊ/.

PETERS & SCHOORMANN 2014). The three durations were determined on the basis of the median durations of short and long vowels obtained by the median speaker. The median duration of the short vowels was 102 ms and the median duration of the long vowels was 193 ms. The duration of the half-long stimuli was approached by the average duration of short and long stimuli, which is 147 ms.

Rising-falling contours were obtained by calculating the f0 values of the beginning and end of the contour as well as the proportional location of the peak in the vowel from the vowel productions of the model speaker. To this end, we determined the median for each of the 10 vowels and calculated the final value as the median of the 10 median values. The proportional location of the peak was at 34% of the total duration of the vowel. To generate the flat contour, we first calculated the average pitch from the f0 values at the beginning and at the peak: $f0_{av\_beginning+peak}$. Second, we calculated the average of the pitch values at the peak and at the end: $f0_{av\_peak+end}$. Finally, the pitch value of the flat contour was calculated as $[0.34 \times f0_{av\_beginning+peak}] + [(1\text{-}0.34) \times f0_{av\_peak+end}]$. Since the location of the peak was at 34% of the vowel duration, the average of the pitch at the beginning and at the peak was weighed by 0.34 in this formula, and the average of the pitch at the peak and at the end was weighed by 1-0.34 = 0.66. The resulting six stimulus versions per vowel are listed in Table 1. The six contours are schematically visualized in Figure 3.

Table 1. Time (ms) and f0 (Hz) values used for creating the six stimulus versions.

Duration	Contour	Beginning		Peak		End	
		Time	f0	Time	f0	Time	f0
short	peak	0	126	34	130	102	91
half-long	peak	0	126	50	130	147	91
long	peak	0	126	65	130	193	91
short	flat	0	116	34	116	102	116
half-long	flat	0	116	50	116	147	116
long	flat	0	116	65	116	193	116

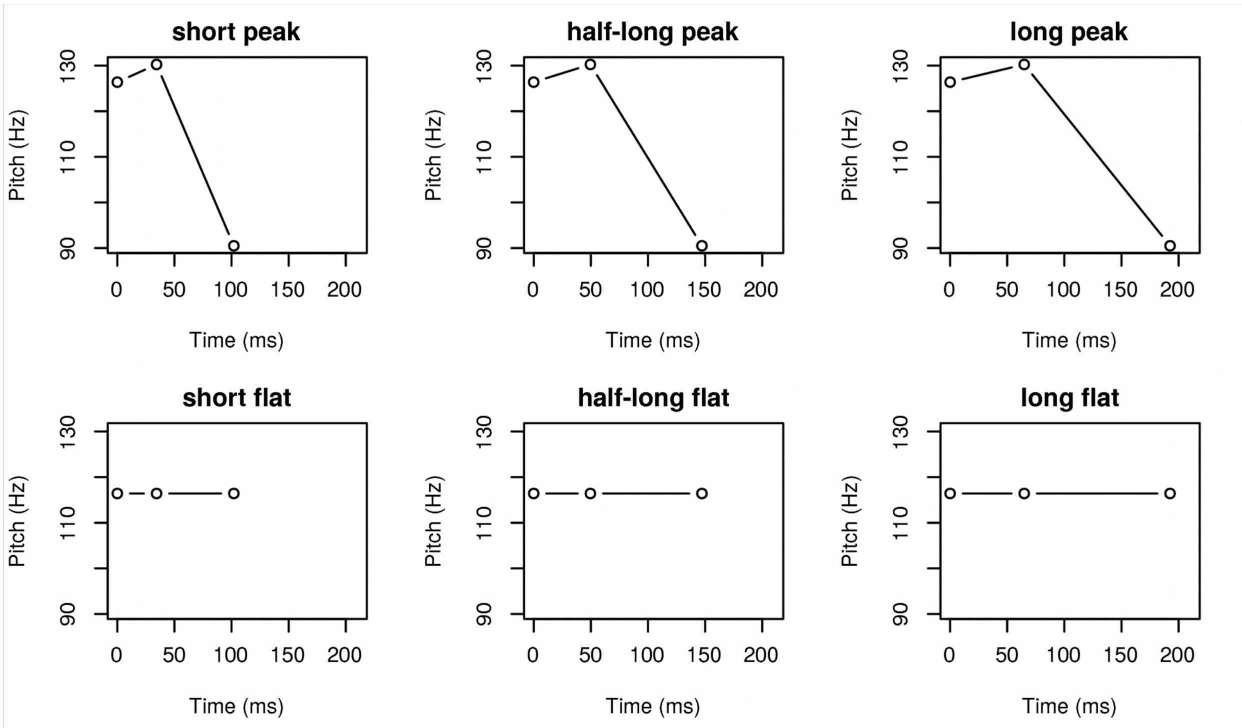

Figure 3. Schematic visualization of the contours of the six stimulus versions per vowel. The peak of the rising-falling contours occurs at 34% of vowel duration.

2.3 Procedure

The vowel stimuli were presented in a perception task. On each screen, the six versions of a vowel were randomly assigned to the letters A to F and embedded as listening samples. A small speaker icon was placed to the right of each letter, as illustrated in Figure 4. When clicking on this icon, the embedded stimulus was played over headphones (Sure SRH2YY). All experiments were conducted as self-paced experiments in a quiet room. Subjects could play each stimulus as often as they wanted. They were instructed to order the stimuli from shortest to longest, based on their perception of syllable duration. A vertical scale from 1 to 6 was provided on a paper, where 1 meant 'shortest duration' and 6 'longest duration'. The subjects were asked to assign each of the stimuli to one of the six levels by writing the letter of the stimulus to the right of the level number. Multiple stimuli were allowed to be assigned to the same level, and not all levels had to be used.

The experiment consisted of three blocks with forced breaks between each block, one 1.5-minute break between the first and second block, and one 3-minute break between the second and third block. Subjects were allowed to extend the breaks as long as they wanted. The first block served as a short training session in which two out of the 10 vowels were presented. The second and third block comprised the perception test including the full vowel set per block. The stimuli were presented in a randomized order so that the order of the stimuli differed per subject and per block.

Figure 4. Example screen of the experiment as presented to the subjects. The screens are numbered from 1 to 22. The present example shows the first screen, as indicated by the number in the upper right corner.

2.4 Statistical analysis

We looked for effects of DURATION (short – half-long – long), CONTOUR (flat – rising-falling), VOWEL QUALITY (iː ɪ uː ʊ eː ɛː ɛ oː ɔː ɔ) and the interaction between DURATION and CONTOUR, VOWEL QUALITY and CONTOUR as well as the three-way interaction of DURATION, CONTOUR, and VOWEL QUALITY. Since the ratings were on an ordinal scale, we used the *clmm* function in the R package *ordinal*, which enabled us to perform a cumulative link mixed model (R Core Team 2015, CHRISTENSEN 2015). Model fit was determined through the AIC value. All models are given in the appendix. We used the function *lsmeans* from the *lsmeans* package for multiple comparisons of factors with three or more levels and the Bonferroni method to adjust the *p*-values.

Note that SEX is included as a control variable in the model due to the participation of both male and female listeners in the experiment. The inclusion of both sexes was a necessity in order to get a sufficient number of trilingual speakers. Accordingly, both sexes were also included among the monolingual speakers to obtain a balanced set. No effects for speaker sex are anticipated for the perception of vowel duration or the f0 cue.

3. Results

3.1 Overall effects of f0 dynamics across the two locations

The effects of vowel DURATION (short, half-long, and long) and CONTOUR (flat or dynamic f0) are shown in Figures 5 and 6 respectively. Table 2 shows the results of mixed-effects models of the Scharrel and Hanover data.

Table 2. Effect of DURATION and CONTOUR for Scharrel and Hanover.

				Scharrel		Hanover	
				z ratio	sig.	*z* ratio	sig.
duration	short	vs.	half-long	-12.626	p<.0001	-16.625	-12.626
	short	vs.	long	-15.960	p<.0001	-15.812	-15.960
	half-long	vs.	long	-15.540	p<.0001	-12.265	-15.540
contour	flat	vs.	peak	-1.194	n.s.	-2.530	-1.194

Subjects from both locations distinguished between short, half-long, and long stimuli (see Figure 5 and Table 2). In addition, the monolingual subjects showed a general effect of a dynamic f0 in the perception of vowel duration. Stimuli with a dynamic f0-contour were rated as longer than the ones with a flat f0 contour (see Figure 6 and see Table 2). No overall lengthening effect was found for the trilingual speakers.

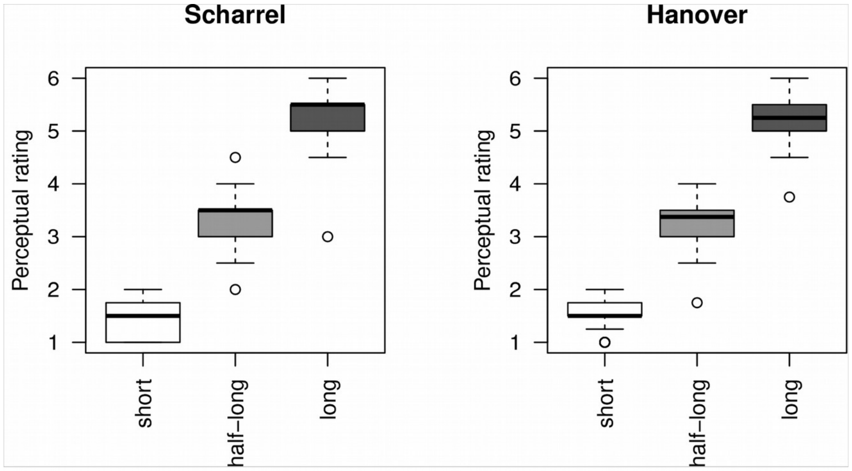

Figure 5. Median perceptual duration ratings of vowels per location and duration.

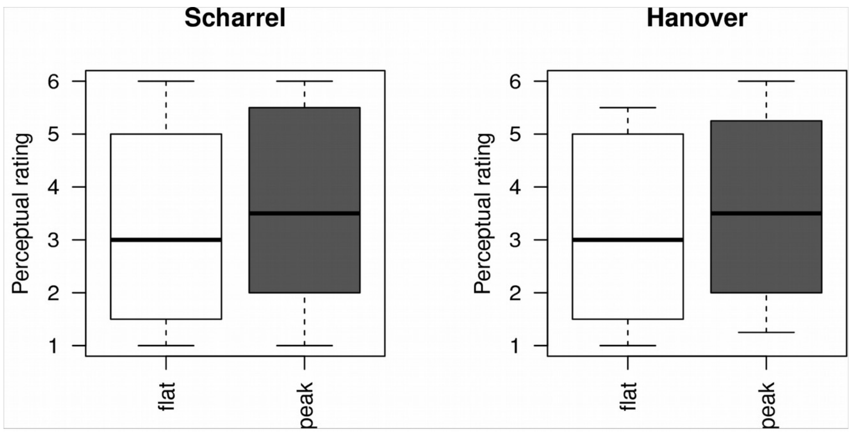

Figure 6. Median perceptual duration ratings of vowels per location and contour.

In addition, the analysis shows a significant difference in the overall perception of vowel duration between male and female speakers from Scharrel. The male speakers generally perceived the duration of all vowels as longer than the females. This effect was not anticipated and cannot be meaningfully interpreted at this point.

3.2 Overall interaction effects of duration and f0 dynamics across the two locations

The effects of the interaction between vowel DURATION and CONTOUR on perceived vowel duration are shown in Figure 7. Table 3 shows the results of mixed-effects models of the Scharrel and Hanover data.

Table 3. Effect of the interaction between DURATION and CONTOUR for Scharrel and Hanover.

				Scharrel		Hanover	
				z ratio	sig.	*z* ratio	sig.
interaction	short flat	vs.	short peak	-0.038	n.s.		n.s.
	half-long flat	vs.	half-long peak	-3.729	p<.01		n.s.
	long flat	vs.	long peak	-8.429	p<.001		n.s.

In Scharrel, only the perceived duration of half-long and long vowels was increased by a dynamic f0 when compared to a flat f0 contour (see Figure 7 and Table 3). For the short vowels no interaction of vowel duration and contour was found. No interaction between vowel duration and contour was found in Hanover ratings, suggesting that there are no differences in the way the contour affected the three vowel durations. Stimuli of each duration with a dynamic f0 contour were always rated as longer than stimuli with a flat f0 contour by the Hanover subjects. Figure 7 illustrates for the Hanover ratings that stimuli with a short vowel segment and a flat contour have about the same median as stimuli with a short vowel and a peak. However, the boxes hardly overlap. For short vowels with a flat contour the median is equal to the third quartile, while the median of vowels which have a rising-falling contour have a median equal to the first quartile.

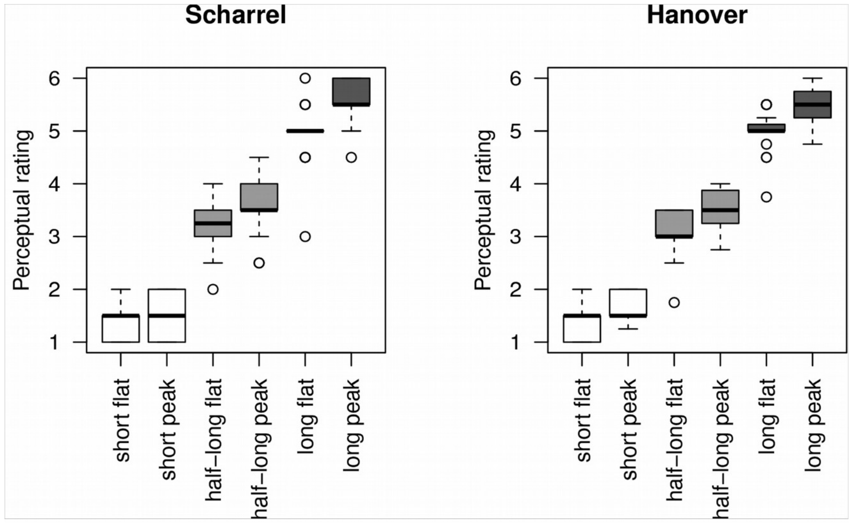

Figure 7. Median perceptual duration ratings of vowels per location, duration, and contour.

3.3 Interaction effects of vowel quality and f0 dynamics across the two locations

We also examined the interaction of VOWEL QUALITY and CONTOUR as well as the three-way interaction of DURATION, CONTOUR, and VOWEL QUALITY to test whether dynamic f0 affects the perceived vowel duration and the three durations for each vowel quality differently. No three-way interaction of DURATION, CONTOUR, and VOWEL QUALITY was found for either the monolingual listeners or the trilingual speakers. This means that for each of the 10 vowel qualities all speakers perceived the effect of a dynamic f0 equally for all three durations. While the trilingual speakers show a general sensitivity to the f0 cue only in the half-long and long vowels when all vowels are considered together (see 3.2), no such concentration on the two degrees of vowel duration is found when individual vowel qualities are considered.

Table 4 and Figure 8 show the results of the mixed-effects models of the Scharrel and Hanover data. The lengthening effect of a dynamic f0 contour was found for half-close tense /oː/ and half-open lax /ɔ/ in all speakers, regardless of language background (see Table 4). In addition, only the monolingual subjects perceived the half-open lax front vowel /ɛ/ as longer when realized with a rising-falling f0 contour.

Table 4. Effect of the interaction between VOWEL QUALITY and CONTOUR for Scharrel and Hanover.

				Scharrel		Hanover	
				z ratio	sig.	*z* ratio	sig.
interaction	/oː/ flat	vs.	/oː/ peak	-5.879	*p*<.0001	-6.139	*p*<.0001
	/ɔ/ flat	vs.	/ɔ/ peak	-4.530	*p*<.01	-5.829	*p*<.0001
	/ɛ/ flat	vs.	/ɛ/ peak	-0.851	n.s.	-3.970	*p*<.05

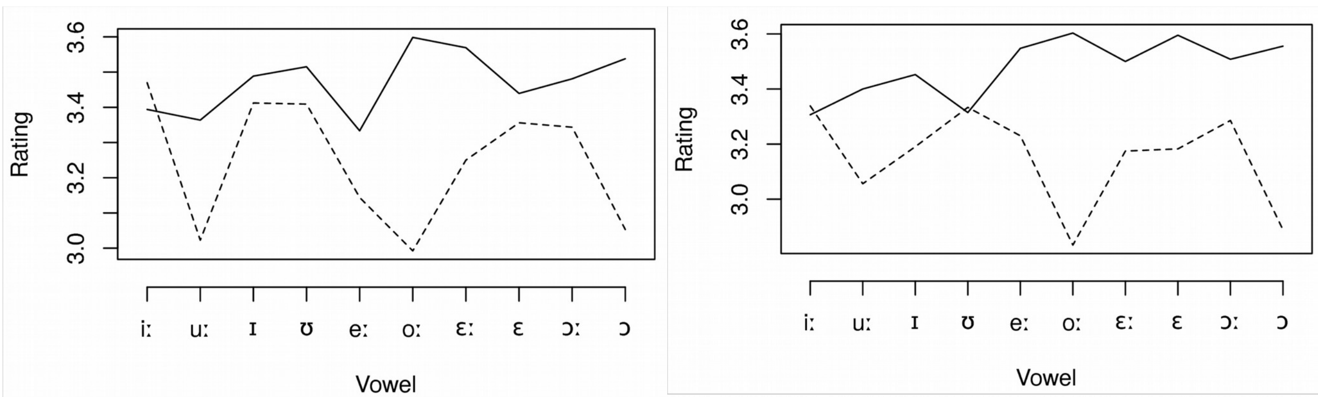

Figure 8. The interaction between VOWEL QUALITY and CONTOUR for Scharrel (left) and Hanover (right). In the two plots the dashed line represents average ratings for vowels with a flat contour, and the solid line represent the average ratings for vowels with a dynamic contour.

Figure 8 suggests that a larger increase of perceived duration is especially found in those vowels which were perceived as relatively short when realized with a flat contour. This observation is supported by a post hoc analysis of the correlation of the stimulus' ratings with and without a dynamic f0. The correlation between the mean duration ratings of vowels with a flat contour and the corresponding mean increase in the perceived duration of vowels with a dynamic contour is $r=-.91$ ($p<.001$) for Scharrel and $r=-.94$ ($p<.0001$) for Hanover.

4. Conclusion and discussion

All subjects distinguished short, half-long, and long stimuli across all vowels irrespective of their language background. Scharrel and Hanover used acoustic vowel duration independently of the f0 contour. This finding concurs with previous notions of vowel duration as a readily accessible and most salient cue to vowel length distinction (BOHN 1995, cf. LEHNERT-LEHOUILLER 2010).

In the present study we were particularly interested in the role that f0 dynamics play in the perception of vowel duration. First, we examined whether vowel stimuli of equal acoustic duration would be perceived as relatively shorter or longer depending on the f0 contour (question 1). We found a difference in the overall perception of the f0 cue between the two speaker groups. Only the monolingual subjects' ratings showed the general lengthening effect of dynamic f0 when all ten monophthongs were analyzed together. This result is in line with VAN DOMMELEN (1991, 1993) who found a perceptual lengthening effect of dynamic f0 in German for isolated monosyllabic words.

Second, we were interested in whether this effect was limited to specific vowel durations (questions 2). Interaction effects showed differences in the ratings for the three durations in the trilingual speakers only. The Hanover monolinguals always perceived stimuli with a rising-falling pattern as longer than stimuli of the same duration with a flat f0 contour irrespective of stimulus duration. In the trilingual Scharrel speakers, the perceptual lengthening was limited to the half-long and long vowel stimuli. This finding agrees with our expectation that the trilingual listeners would show a specific use of this feature rather than a general sensitivity towards dynamic f0.

To further study the specific use of the dynamic f0 contour, we examined the perceived lengthening effect for the individual vowel categories (question 3). There was a general lengthening effect of the dynamic f0 contour for /ɔ oː/ in both speaker groups and additionally for /ɛ/ in the monolingual speaker group. The expected increased sensitivity of the trilinguals regarding close tense and possibly open lax vowels, however, is not supported by our data. It can be argued that since the close tense vowels are merged in present-day Saterland Frisian (SCHOORMANN, HEERINGA & PETERS

2017a), the distinction between short and long close tense vowels and with it the use of a dynamic f0 contour as a secondary feature to vowel length distinction as described in HEERINGA, PETERS & SCHOORMANN (2014) is lost in most present-day speakers. Hence, the outcome of the perception task does not suggest a language-specific use of the f0 cue that can be reduced to differences in the vowel inventories.

The question remains why a lengthening effect was found for only three vowel categories. We found that the increase in perceived duration was larger for vowels that are perceived as relatively short when realized with a flat contour. Since the durations of vowels do not differ across all categories for each of the three durations (i.e. short, half-long or long), we suggest that the variation in the perceived duration of vowels with a flat contour is the result of variation in spectral properties. This idea, however, requires a further targeted investigation.

However, the trilinguals use dynamic f0 to increase the contrast between short vowels versus half-long/long vowels. If short vowels do not sound longer as a result of f0 dynamics but half-long/long vowels do, the contrast is increased. For the monolinguals of Northern High German, however, no such effect was found. Longer vowels sound longer, but if half-long and short vowels do as well, the mutual contrasts remain the same. This observation can be applied to the differences in acoustic vowel duration between monolingual speakers from Hanover and trilingual speakers from Scharrel, which were observed in a previous production experiment by SCHOORMANN, HEERINGA & PETERS (2017b) who showed that the duration ratio of short and long oppositions is considerably smaller in the trilingual speakers than in the monolingual speakers. The smaller ratios in all of the trilinguals' three languages were mainly due to the shorter acoustic durations of the long vowels. We argue that the trilingual speakers use the f0 cue as a secondary acoustic feature to enhance the contrast between short and long oppositions.

Our results confirm a general effect of dynamic f0 on perceived vowel duration irrespective of language background. The findings are therefore in agreement with previous studies in as far as they suggest that the perception of vowel duration may be influenced by f0 dynamics irrespective of the subjects' language background. Furthermore, they are in line with the notion that the specific use of f0 as a secondary cue is likely to be language-specific (cf. LEHNERT-LEHOUILLIER 2010). These findings add to a body of research on the perceptual lengthening effect and its language-specific use, especially with regard to languages with a large vowel inventory. The relation between the lengthening effect and the language background of the speaker certainly deserves further consideration.

References

Bohn, Ocke-Schwen / James Emil Flege (1992): The production of new and similar vowels by adult German learners of English. In: Studies in Second Language Acquisition 14, 131–158.

Bohn, O.-S. (1995). Cross language speech production in adults: First language transfer doesn't tell it all. In Strange, Wilfried (Ed.), Speech perception and linguistic experience: Issues in cross-language research. Baltimore: York Press, 279–304

Bohn, Ocke-Schwen (2004): How to organize a fairly large vowel inventory. The vowels of Fering (North Frisian). In: Journal of the International Phonetic Association 34, 161–173.

Christensen, Rune Haubo Bojesen (2015): ordinal – Regression Models for Ordinal Data. [R package, Version 2015.6-28, https://CRAN.R-project.org/package=ordinal].

Cumming, Ruth (2011): The effect of dynamic fundamental frequency on the perception of duration. In: Journal of Phonetics 39, 375–387.

Elmentaler, Michael (2012): In Hannover wird das beste Hochdeutsch gesprochen. In Anderwald, Lieselotte (Ed.), Sprachmythen – Fiktion oder Wirklichkeit. Frankfurt/M.: Peter Lang, (Kieler Forschungen zur Sprachwissenschaft 3),101-115.

Fort, Marron C. (2004): Sprachkontakt im dreisprachigen Saterland. In: Munske, Horst Haider (Hg.): Deutsch im Kontakt mit germanischen Sprachen. Tübingen: Niemeyer (Reihe Germanistische Linguistik. 248), 77–98.

Fort, Marron Curtis (2015): Saterfriesisches Wörterbuch mit einer phonologischen und grammatischen Übersicht. Hamburg: Buske.

Heeringa, Wilbert / Jörg Peters / Heike Schoormann (2014): Segmental and prosodic cues to vowel identification. The case of /ɪ i iː/ and /ʊ u uː/ in Saterland Frisian. In: Proceedings of the 7th International Conference on Speech Prosody, Dublin, May 20–23, 2014, 643–647.

Heeringa, Wilbert / Heike Schoormann / Jörg Peters (2015): Cross-linguistic vowel variation in Saterland: Saterland Frisian, Low German, and High German. In: Proceedings of the 18th

International Congress of Phonetic Sciences, Glasgow, 10–14 Aug. 2015, paper number 0443.

Jørgensen, Hans Peter (1969): Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantenfrequenzen. Phonetica 19, 217–245.

Kinoshita, Keisuke / Dawn Marie Behne / Takayuki Arai (2002): Duration and F0 as perceptual cues to Japanese vowel quantity. In: Proceedings of the 7th International Conference on Spoken Language Processing, Denver, September 16-20, 2002, 757–760.

Kohler, Klaus Jürgen (1995): Einführung in die Phonetik des Deutschen (Grundlagen der Germanistik. 20). Berlin: Erich Schmidt.

Kramer, Piet (1982): Kute Seelter Sproakleere. Rhauderfehn: Ostendorp.

Lehiste, Ilse (1976): Influence of fundamental frequency pattern on the perception of duration. In: Journal of Phonetics 8, 469–474.

Lehnert-LeHouillier, Heike (2010): A cross-linguistic investigation of cues to vowel length perception. In: Journal of Phonetics 38, 472–482.

Lippus, Pärtel / Karl Pajusalu / Jüri Allik (2011): The role of pitch cue in the perception of the Estonian long quantity. In: Frota, Sónia / Gorka Elordieta / Pilar Prieto (Hg.): Prosodic categories. Production, perception and comprehension. Dordrecht: Springer (Studies in Natural Language and Linguistic Theory. 82), 231–242.

Pätzold, Matthias / Adria P. Simpson (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. In: Adrian P. Simpson, Klaus J. Kohler, & Tobias Rettstadt (Hg.): The Kiel Corpus of Read/Spontaneous Speech - Acoustic data base, processing tools and analysis results. Kiel: Institut für Phonetik und digitale Sprachverarbeitung (Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK). 32), 215-247.

Peters, Jörg (2008): Saterfrisian intonation. An analysis of historical recordings. In: Us Wurk 57, 141–169.

Pisoni, David B. (1976): Fundamental frequency and perceived vowel duration. In: Journal of the Acoustical Society of America 59, S39.

R Core Team (2015): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [https://www.R-project.org/].

Siebs, Theodor (1966 [1889]): Zur Geschichte der englisch-friesischen Sprache. Reprint. Wiesbaden: Sändig [Halle a. d. Saale: Niemeyer].

Schoormann, Heike / Wilbert Heeringa / Jörg Peters (2017a): A cross-dialectal acoustic study of Saterland Frisian vowels. In: Journal of the Acoustical Society of America 141 (4), 2893-2908; doi: http://dx.doi.org/10.1121/1.4980855.

Schoormann, Heike / Wilbert Heeringa / Jörg Peters (2017b): Standard German vowel productions by monolingual and trilingual speakers. In: International Journal of Bilingualism, doi: https://doi.org/10.1177/1367006917711593, first published online June 27, 2017.

Šimko, Juraj / Daniel Aalto / Pärtel Lippus / Marcin Włodarczak / Martti Vainio (2015): Pitch, perceived duration and auditory biases. Comparison among languages. In: Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, 10–14 Aug. 2015, paper number 0575.

Sjölin, Bo (1969): Einführung in das Friesische. Metzler: Stuttgart.

Steinlen, Anja (2005): The influence of consonants on native and non-native vowel production. A cross-linguistic study. Tübingen: Gunter Narr (Tübinger Beiträge zur Linguistik Ser. A Language development. 30).

Stellmacher, Dieter (1998): Das Saterland und das Saterländische. Oldenburg: Isensee (Vorträge der Oldenburgischen Landschaft. 30).

Tröster-Mutz, Stefan (1997): Phonologie des Saterfriesischen. Überarb. Vers. der Magisterarbeit 1995, Universität Osnabrück, FB Sprach- und Literaturwissenschaften.

Tröster-Mutz, Stefan (2002): Untersuchungen zu Silbenschnitt und Vokallänge im Saterfriesischen.

Theorie des Lexikons (Arbeiten des SFB 282. 120), 1–27.

van Dommelen, Wim A. (1991): F0 and the perception of duration. In: Proceedings of the 12th International Congress of Phonetic Sciences, Aix-en-Provence, August 19-24, 1991, 2:282–285

van Dommelen, Wim A. (1993): Does dynamic F0 increase perceived duration? New light on an old issue. In: Journal of Phonetics 21, 367–386.

Wang, William S.-Y. / Ilse Lehiste / Chin-Kuang Chuang / Nancy Darnovsky (1976): Perception of vowel duration. In: Journal of the Acoustical Society of America, 60, S92.

Wiese, Richard (2000): The phonology of German. Oxford: Oxford University Press.

Yu, Alan C. L. (2010): Tonal effects on perceived vowel duration. In: Laboratory phonology 10, 151–168.

Appendix: Ordinal mixed-effects models used for explaining perceived vowel duration

Final Model 1 (Scharrel):

response

rating~

fixed effects

sex + duration + contour + vowel_quality + sex:duration +

sex:contour + duration:contour + contour:vowel_quality +

random effects

(1 + duration | informant)

Final Model 2 (Hanover):

response

rating~

fixed effects

duration + contour + vowel_quality + contour:vowel_quality +

random effects

(1 + duration + contour | informant)