Segmental and prosodic cues to vowel identification: The case of /ɪ i iː/ and /ʊ u uː/ in Saterland Frisian

Wilbert Heeringa[1], *Jörg Peters*[1], *Heike Schoormann*[1]

[1]Institute of German Studies, Carl von Ossietzky University, Oldenburg, Germany

wilbert.heeringa@uni-oldenburg.de, joerg.peters@uni-oldenburg.de, heike.schoormann@uni-oldenburg.de

Abstract

Saterland Frisian has a complete set of closed short tense vowels. Together with the long tense vowels and the short lax vowels they constitute series of phonemes that differ by length and/or tenseness. We examined the cues that distinguish the front unrounded and the back rounded series of short lax and short and long tense vowels in triplets by eliciting 'normal speech' and 'clear speech' in a reading task from two speakers. Short and long vowels were distinguished by vowel duration, and lax and tense vowels by their location in the F1-F2 space. The durational difference between short tense and long tense vowels, however, was largely restricted to the 'clear speech' condition. In 'clear speech', 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 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.

Index Terms: f0 dynamics, f0 excursion, formants, Saterland Frisian, tenseness, vowel duration

1. Introduction

Saterland Frisian is spoken in three small villages – Strücklingen, Ramsloh and Scharrel – in the north-western corner of the district of Cloppenburg in Lower Saxony. It is the only remaining living variety of Old East Frisian, which was spoken along the coasts of the Netherlands and Lower Saxony. Saterland is believed to have been colonised by Frisians from the coastal areas in the eleventh century. According to the most recent count, Saterland Frisian is spoken by 2250 speakers [1].

Saterland Frisian has a complete set of closed short tense vowels: /i y u/ [2], [3], [4]. 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. Potential acoustic cues which distinguish the vowels in a triplet may be vowel duration, spectral features (F1, F2) as well as the timing and scaling of f0.

In this paper we investigate which acoustic cues distinguish the sounds within two triplets containing /ɪ i iː/ and /ʊ u uː/ respectively. For each of the two triplets we conducted a traditional reading task and a listener-directed task, which maximizes the discrimination between words and is an effective way to reveal potential segmental and prosodic cues.

Several studies show that vowels with stronger f0 dynamics are perceived as being longer. Lehiste [5] found that listeners perceived a falling-rising or rising-falling f0, as opposed to a flat f0 pattern, to be longer even when the stimuli have the same acoustic duration. Yu [6] found the same effect for dynamic versus flat f0 and also showed that syllables with higher f0 are heard as longer than syllables with lower f0. Cumming [7] was able to show the perceived lengthening effect of dynamic f0 for native speakers of Swiss German, Swiss French and French. This effect is likely language-specific [8].

The acoustic cues, which add to the distinction of vowels in Saterland Frisian triplets have not yet been fully studied. Siebs [9] distinguishes between tone accents in Saterland Frisian (*Stoßton* versus *Schleifton*) which suggests that f0 might play a role. In a more recent study Tröster-Mutz [10] [11] investigated the phonetics of Saterland Frisian vowels, but did not find any evidence for tone accent differences in present-day Saterland Frisian. In this paper, we focus on the question whether f0 dynamics have any systematic effect which may help to discriminate between the vowels of each triplet.

2. Method

2.1. Material

The two triplets used are still known by a restricted number of Saterland Frisian speakers. For the closed front vowels /ɪ i iː/ we used *Smitte* 'forge', *smiete* 'to throw' and *Smíete* 'throws' (pl.). The closed back vowels /ʊ u uː/ were elicited by the triplet *ful* 'full', *fuul* 'rotten' and *fúul* 'much'.

2.2. Procedure

For each of the triplets we conducted two experiments, one eliciting 'normal speech' and another eliciting 'clear speech'. The experiments were carried out by two female native speakers, aged 78 and 66 years, henceforth referred to as subject 1 and subject 2 respectively. The two speakers are born and raised in Ramsloh and have lived in this village most of their lives. We chose Ramsloh since it is located in the center of Saterland and its Saterland Frisian variety is considered to be the most conservative [3].

2.2.1. Normal speech

In this experiment Saterland Frisian words were presented in written form to the two native speakers on a computer screen, one word at a time. We used twelve different words: six triplet words (*ful*, *fuul*, *fúul*, *Smitte*, *smiete*, *Smíete*), and six filler words (*Pot*, *Paad*, *Kat*, *leet*, *Täk*, *Poot*).

A session consisted of four blocks in which each of the 12 words was presented four times. Within each block the order of the words was randomized, so that a word was never followed by the same word or by a word belonging to the same triplet. Three of the six filler words (*Pot*, *Paad* and *Kat*) were also used as a short practice, preceding the first block. In sum, 195 words were presented in one session.

We obtained 16 samples per subject per triplet word. Looking for cues concerning the f0 dynamics, only samples with a clear f0 peak in the vowel can be used for the analysis. The number of word samples which satisfy this condition is given per triplet word and per subject in Table 1.

2.2.2. *Clear speech*

Saterland Frisian words were presented in written form to the two native speakers on a computer screen. In this condition, only the six triplet words were used. For maximum discrimination, a triplet word was always presented together with the two other triplet words. The word to be pronounced was encircled and displayed in blue (the other words were black). The three words of a triplet were located on the screen so that they were imaginary vertices of a triangle. Each 'triangle' was rotated over an arbitrary angle.

One session consisted of four blocks. Each of the triplet words was presented eight times per block. Within a block, 24 words of the /ʊ u uː/ triplet were presented first, followed by 24 words of the /ɪ i iː/ triplet. Thus, in each block 48 words were pronounced. In each part the words were presented in a randomized order so that a word was not followed by the same word.

In this experiment, the subjects were either speaker or listener and changed roles after each block. When one subject read the words aloud, the other subject marked the triplet word she thought she heard. The reader and the listener were separated by a screen during the experiment.

We obtained 16 samples per speaker per triplet word. Just as for 'normal speech' we limited the analysis to word samples with a clear f0 peak when looking for cues concerning the f0 dynamics. The number of word samples with f0 peak are given per triplet word and per speaker in Table 1.

Table 1: *Number of word samples with a clear f0 peak per triplet word and per subject.*

	Normal speech		Clear speech	
	sub1	sub2	sub1	sub2
Smitte	14	16	15	14
smiete	15	16	16	15
Smíete	13	16	16	16
ful	1	13	6	5
fuul	10	16	5	11
fúul	7	15	15	10

2.3. Acoustic variables

Segmental and prosodic variables were measured with PRAAT [12]. For each word belonging to the /ɪ i iː/ triplet we measured the duration of each of the segments: /s/, /m/, V, /t/ and the final schwa. The duration of /t/ was split into two parts: the time from the beginning of the segment to the burst (t1), and from the burst to the end of the segment (t2). We also measured spectral variables F1 and F2 at the center of the vowel.

We measured f0 variation in the interval from the beginning of /m/ to the end of V. When the f0 peak was somewhere in this interval, we considered a rise and a fall. F0 dynamics was operationalized by calculating the relative f0 excursion in semitones per millisecond, which was obtained by calculating the sum of the f0 rise size (pitch of f0 peak minus pitch at the beginning of the interval) and the f0 fall size (pitch of f0 peak minus pitch at the end of the interval) divided by the duration of the interval. A related approach was used by Grabe [13].

For each word belonging to the /ʊ u uː/ triplet we likewise measured the duration of each of the segments, /f/, V, and /l/, and F1 and F2 at the center of the vowel. When measuring relative f0 excursion, we focused on the interval starting at the beginning of V and ending at the end of /l/.

2.4. Statistical processing

We looked for acoustic cues that distinguish all of the sounds within a triplet. Per acoustic variable we used a Generalized Linear Model (GLM) where the stimulus was the independent variable and the acoustic variable the dependent variable. Since not all of the acoustic variables are normally distributed, GLM was the fitting test. The stimuli were the triplet words, therefore the stimulus variable was a three level factor. Two-tailed p-values obtained from pairwise comparisons were adjusted using the Bonferroni correction.

The variables were initially analyzed per subject and per triplet. However, in this paper the results of the subjects are combined by showing the consensus. For example, when we found $p < 0.01$ for the one subject, and $p < 0.001$ for the other subject, then the consensus is considered to be $p < 0.01$. When we do not find a significant effect for both of the subjects for a particular variable or a significant difference in opposite direction, no significance is reported. In the tables below, levels of significance are indicated by asterisks: *< 0.05, **< 0.01 and ***< 0.001.

3. Results

3.1. Normal speech

3.1.1. *Results /ɪ i iː/ triplet*

Duration values for the /ɪ i iː/ triplet are shown in the upper panel of Figure 1. Vowel plots are found in Figure 2 and f0 contour plots in Figure 3. Statistical results are summarized in Table 2. *Smitte* is distinguished from *Smíete* by a shorter vowel duration, a higher F1, a lower F2, and smaller f0 dynamics. *Smitte* differs from *smiete* by a shorter vowel duration, a higher F1, a lower F2, a smaller fall size and smaller f0 dynamics. The triplet word *smiete* is distinguished from *Smíete* by larger f0 dynamics. Hence, tense short vowels have largest f0 dynamics.

3.1.2. *Results /ʊ u uː/ triplet*

Duration values for the /ʊ u uː/ triplet are shown in the lower panel of Figure 1. Vowel plots are found in Figure 2 and f0 contour plots in Figure 4. Statistical results are summarized in Table 3. In Table 1 we find that *ful* is represented by just one sample in the normal speech experiment. Therefore comparisons between *ful* and *fuul* and between *ful* and *fúul* are based on subject 2 only. The triplet word *ful* has a smaller vowel duration and higher F1 and F2 values than the other triplet words. Additionally, *ful* has a smaller /f/ duration then *fúul*. The triplet words *fuul* and *fúul* are not distinguished, i.e. short tense and long tense vowels are not distinguished.

3.2. Clear speech

3.2.1. *Results /ɪ i iː/ triplet*

Duration values for the /ɪ i iː/ triplet are shown in the upper panel of Figure 1. The clear speech graph shows a stronger contrast between short vowels and long tense vowels than the normal speech graph. The vowel plots are found in Figure 2. Unlike in normal speech, short tense and long tense front vowels are separated in the acoustic vowel space in clear speech. Figure 3 shows that the difference between rise size and fall size in clear speech has decreased compared to normal speech. Statistical results are summarized in Table 2. The number of variables that distinguish triplet words is much larger in clear speech than in normal speech. When comparing *Smitte* and *Smíete* we find smaller values for the duration /m/, V, t2 and /ə/, a higher F1, a lower F2 and a smaller rise size. We get the same findings when comparing *smiete* and *Smíete*, but rise size is not significant and *smiete* has larger f0 dynamics than *Smíete*. *Smitte* has a smaller vowel duration, a higher F1, a lower F2, a smaller fall size and smaller f0 dynamics than *smiete*. V duration, F1 and F2 distinguish all of the triplet words.

3.2.2. *Results /ʊ u uː/ triplet*

Duration values for the /ʊ u uː/ triplet are shown in the lower panel of Figure 1. Again, the clear speech graph shows a stronger contrast between short vowels and long tense vowels. The vowel plots are given in Figure 2. Eventhough there is still some overlap in the acoustic vowel space for the long and short tense back vowels, the spread and the mean formant values show a clearer separation for the two sounds compared to the normal speech conditon. The f0 contour plots in Figure 4 show a longer f0 fall duration for the long tense vowel compared to the two other vowels. Statistical results are presented in Table 3. *ful* is distinguished from *fúul* by a smaller vowel duration and a higher F1 and F2. *fuul* has a smaller vowel duration, a higher F2, and larger f0 dynamics than *fúul*. *ful* is distinguished from *fuul* by a higher F1.

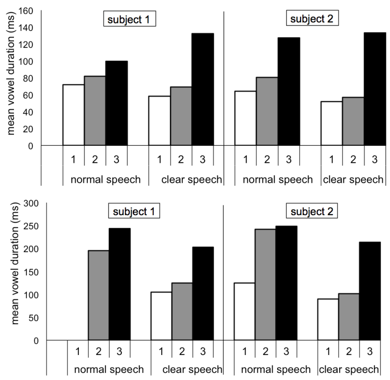

Figure 1: *Duration values for normal speech and clear speech. On top the /ɪ i iː/ triplet (1=Smitte, 2=smiete, 3=Smíete) and at the bottom the /ʊ u uː/ triplet (1=ful, 2=fuul, 3=fúul).*

3.3. Consensus of triplets

In Table 4 we show consensus results of the two triplets. For normal speech we find that lax vowels are distinguished from tense vowels by a higher F1 and a lower (/ɪ i iː/ triplet) or higher (/ʊ u uː/ triplet) F2. Additionally, short lax vowels are distinguished from long tense vowels by a smaller vowel duration. Short tense and long tense vowels are not distinguished by any variable in the normal speech data.

We find that lax and tense short vowels are distinguished only by a higher F1 in clear speech. Short lax vowels are

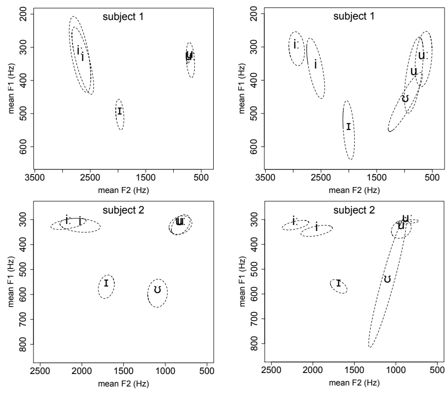

Figure 2: *Vowel plots show the mean formant values of the six triplet sounds for normal speech (left) and clear speech (right). Ellipses enclose two standard deviations.*

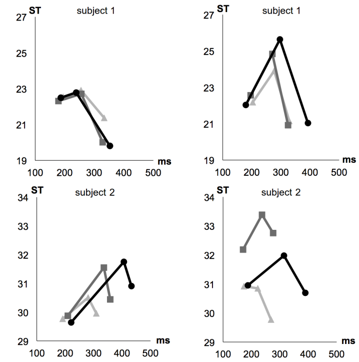

Figure 3: *f0 contours for the /ɪ i iː/ triplet in normal speech (left) and clear speech (right) for subject 1 (top) and subject 2 (bottom). Lighter gray lines with triangles represent short lax vowels, darker gray lines with squares represent short tense vowels, and black lines with circles represent long tense vowels.*

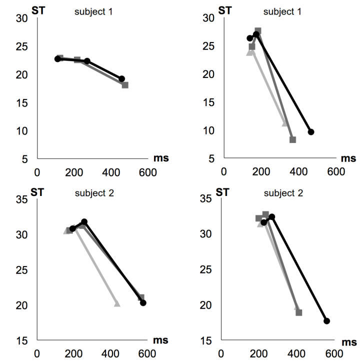

Figure 4: *f0 contours for the /ʊ u uː/. For graphical conventions see Figure 3.*

distinguished from long tense vowels by the same variables as for normal speech. In contrast to normal speech, short tense vowels and long tense vowels are clearly distinguished. Short tense vowels have a smaller vowel duration, a lower (/ɪ i iː/ triplet) or higher (/ʊ u uː/ triplet) F2, and larger f0 dynamics than long tense vowels.

Table 5 shows percentages of stimuli that were correctly predicted on the basis of V duration, F1, F2, and f0 dynamics, and which were obtained by Linear Discriminant Analysis. As expected, percentages for clear speech are higher than for

Table 2: *Results for the /ɪ i iː/ triplet. 1=Smitte, 2=smiete, 3=Smíete.*

	Normal speech			Clear speech		
	sig. 1·2	sig. 1·3	sig. 2·3	sig. 1·2	sig. 1·3	sig. 2·3
/s/ duration						
/m/ duration					**	***
V duration		**		*	***	***
t1 duration						
t2 duration					**	**
/ə/ duration					**	***
F1	***	***		***	***	***
F2	***	***		***	***	***
f0 rise size					**	
f0 fall size	**			*		
f0 dynamics	*	*	*	**		*

Table 3: *Results for the /ʊ u uː/ triplet. 1=ful, 2=fuul, 3=fúul.*

	Normal speech			Clear speech		
	sig. 1·2	sig. 1·3	sig. 2·3	sig. 1·2	sig. 1·3	sig. 2·3
/f/ duration		*				
V duration	***	***			***	***
/l/ duration						
F1	***	***		**	***	
F2	***	***			***	*
f0 rise size						
f0 fall size						
f0 dynamics						**

normal speech. In the case of /ɪ i iː/, all members of the triplet can be distinguished by using vowel duration, F1, and F2 as acoustic cues.

Table 4: *Consensus results of subjects and triplets. For normal speech and clear speech those variables are listed which play a role for both subjects and both triplets. 1=short lax vowel, 2=short tense vowel, 3=long tense vowel.*

speech type	pair	V dur	F1	F2	f0 dyn.
normal	1 · 2		***	***	
	1 · 3	**	***	***	
	2 · 3				
clear	1 · 2		**		
	1 · 3	***	***	***	
	2 · 3	***		*	*

4. Conclusions

We found that vowel duration, spectral variables, and f0 dynamics are cues for the distinction of vowels in two Saterland Frisian word triplets.

Vowel duration. From Table 4 we conclude that vowel duration implements the phonological feature [± long]. In normal speech short lax and long tense vowels are distinguished by vowel duration, whereas in clear speech short tense and long tense vowels are also distinguished, suggesting a division in short and long vowels.

F1 and F2. Table 4 suggests that the feature [± tense] is implemented by spectral features. Lax vowels have a higher F1 and a lower (/ɪ i iː/ triplet) or higher (/ʊ u uː/ triplet) F2 than tense vowels, i.e. lax vowels are more centered than tense vowels (cf. Figure 2). Gussenhoven [14] observed that closed vowels are perceived as relatively longer than open vowels, which suggests that lowering of F1 may increase perceived vowel duration, which can be explained by a tendency of listeners to compensate for the intrinsically shorter duration of closed vowels. In view of this finding, the slight additional centralization of the short tense vowels (/i/ and /u/ cf. Figure 2) relative to /iː/ and /uː/ may be used to enhance the perceived durational contrast between the short and long tense vowels.

Table 5: *Discriminant percentages per triplet word and per subject. For each subject important variables contributing to discrimination are marked by 1 and/or 2.*

triplet	sp. type	V dur	F1	F2	f0 dyn.	sub1 %	sub2 %
/ɪ i iː/	norm.	1,2	1,2	1,2	2	90.5	91.7
	clear	1,2	1,2	1,2		100	100
/ʊ u uː/	norm.		1,2	1	2	55.6	65.9
	clear	1,2	1,2	1,2	1	92.3	96.2

F0 dynamics. Short tense vowels were found to have larger f0 dynamics than long tense vowels and, in case of [ɪ], short lax vowels (cf. Figures 3 and 4). According to [5] and [7], increased f0 dynamics may enhance perceived vowel duration. As shortening of /i/ and /u/ in clear speech resulted in vowels that were hardly longer than the short lax vowels [ɪ] and [ʊ], increased f0 dynamics may increase the perceived durational difference between short lax and tense vowels, at least in the case of [ɪ] and [i]. We found that f0 dynamics have a systematic effect which contributes to the tripartite vowel contrast. The most likely interpretation of this contribution is phonetic feature enhancement. The possibility that variation of tonal structure is involved as suggested by Siebs' terms *Stoßton* and *Schleifton* [9] cannot be determined on the basis of our current data, but we will investigate this in a later study.

Overall, our data suggest that the phonological contrasts between [ɪ i iː] and between [ʊ u uː] can be accounted for by the combination of two distinctive features, [± long] and [± tense]. The slight centralization of /i/ and /u/ relative to /iː/ and /uː/ in clear speech may be used to enhance the durational contrasts between the short and long tense vowels. Increased f0 dynamics of tense short vowels relative to lax short vowels, which was found for [ɪ] and [i], may be regarded as a means to make the short tense vowels sound more different from the short lax vowels.

Kohler [15] investigated triplets of closed vowels in High German and in Low German dialects spoken in Schleswig-Holstein and Lower Saxony. In some dialects he found that lax and tense vowels within a triplet differ qualitatively, and short and long tense vowels differ quantitatively. Just as Kohler we found that short lax and tense vowels are distinguished by spectral features only. The clear speech experiment revealed that short and long tense vowels are dinstinguished by vowel duration, F2, and f0 dynamics, being a combination of qualitative and quantitative variables.

5. Acknowledgements

We would like to thank the two Saterland Frisian informants for participating in our experiments. We are grateful to Darja Appelganz and Nicole Mayer for labeling the recordings in PRAAT. The research reported in this paper has been funded by the *Deutsche Forschungsgemeinschaft* (DFG), grant number PE 793/2-1.

6. References

[1] Stellmacher, D., “Das Saterland und das Saterländische”, Florian Isensee, 1998.

[2] Sjölin, B., “Einführung in das Friesische”, Metzler, 1969.

[3] Fort, M.C., “Saterfriesisches Wörterbuch mit einer grammatischen Übersicht”, Buske, 1980.

[4] Kramer, P., “Kute Seelter Sproakleere – Kurze Grammatik des Saterfriesischen”, Ostendorp, 1982.

[5] Lehiste, I., “Influence of fundamental frequency pattern on the perception of duration”, Journal of Phonetics 4, 113–117, 1976.

[6] Yu, A. C. L., “Tonal effects on perceived vowel duration”, in C. Fougeron, B., Kuehnert, M., D’Imperio, M. and N. Vallée [Eds], Laboratory Phonology 10, 151–168, Mouton de Gruyter, 2010.

[7] Cumming, R. E., “The effects of dynamic fundamental frequency on the perception of duration”, Journal of Phonetics 39, 375–387, 2011.

[8] Lehnert-LeHouillier, H., “A cross-linguistic investigation of cues to vowel length perception”, Journal of Phonetics 38, 472–482, 2010.

[9] Siebs, Th., “Zur Geschichte der englisch-friesischen Sprache”, Max Niemeyer, 1889.

[10] Tröster-Mutz, S., “Phonologie des Saterfriesischen”, MA Thesis University of Osnabrück, 1997.

[11] Tröster-Mutz, S., “Untersuchungen zu Silbenschnitt und Vokallänge im Saterfriesischen”, Theorie des Lexikons 120, Heinrich-Heine-University, Düsseldorf, 2002.

[12] Boersma, P. and Weenink, D., “Praat: Doing Phonetics by Computer”, available at http://www.praat.org, 1992-2013.

[13] Grabe, E., “Pitch accent realization in English and German”, Journal of Phonetics 26, 129–143, 1998.

[14] Gussenhoven, C., “Vowel height split explained: Compensatory listening and speaker control”, in J. Cole and J.I. Hualde [Ed], Laboratory Phonology 9, 145–172, Mouton de Gruyter, 2007.

[15] Kohler, K.J., “Überlänge im Niederdeutschen?”, in R. Peters, H.P. Pütz and U. Weber [Eds], Vulpis Adolatio. Festschrift für Hubertus Menke zum 60. Geburtstag, 385–402. C. Winter, 2001.