US WURK LXVI (2017)

⌊1382⌋

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

Wilbert Heeringa*, Heike Schoormann**, Jörg Peters**

Abstract. *Saterland Frisian is spoken in Saterland in northwest Germany. This language 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. We investigated which acoustic cues distinguish the sounds within two triplets containing /ɪ i iː/ and /ʊ u uː/ respectively, conducting a traditional reading task in order to obtain 'normal speech' and a listener-directed task in order to obtain 'clear speech'. In the normal speech condition we found for both triplets that short lax and tense vowels were distinguished by F1 and F2. Short and long tense vowels within the /ɪ i iː/ triplet were distinguished by vowel duration and F2. For the /ʊ u uː/ triplet we did not find any cue that distinguishes short and long tense vowels. However, in clear speech, we found that short and long tense vowels in the /ʊ u uː/ triplet are distinguished by vowel duration. In normal speech and clear speech short lax and long tense vowels are furthermore distinguished by f0 fall size. In clear speech we find for both triplets that short tense and long tense vowels are distinguished by f0 dynamics which was calculated by dividing 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) by the duration of the interval. In an additional perception experiment we found a strong agreement between the intended pronunciation of the triplet words and their perception with F2 serving as the best predictor.*

Keywords. *f0 dynamics, clear speech, Frisian, Saterland Frisian, tenseness, vowel duration*

Introduction

Saterland Frisian is spoken in the three villages Scharrel, Ramsloh, and Strücklingen in the north-western corner of the district of Cloppenburg in Lower Saxony in Germany (Fig. 1). It is the only remaining living variety of

Us Wurk, jiergong 66 (2017), s. 27-76

Old East Frisian, which was spoken along the coasts of the Netherlands and Lower Saxony between the rivers Lauwers and Weser. According to the most recent count, Saterland Frisian is spoken by about 2250 speakers (Stellmacher, 1998). In Ramsloh, 40% of the population speaks Saterland Frisian, in Scharrel 29%, and in Strücklingen 26%. In Sedelsberg, which historically does not belong to the Saterland, less than 10% of the population speaks Saterland Frisian.

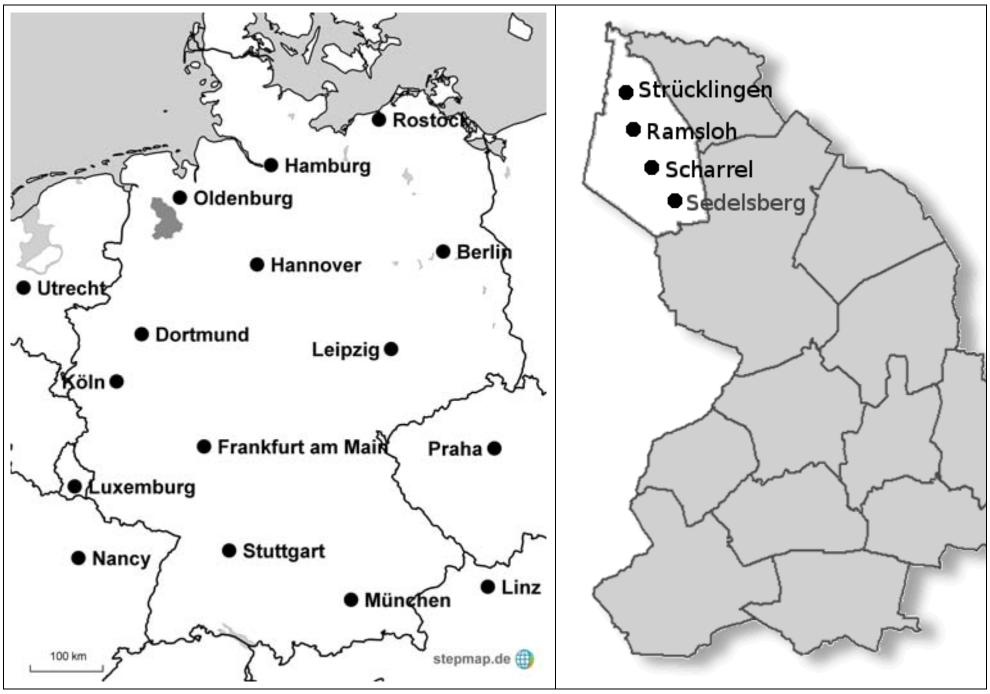


Figure 1. *The Saterland is found in the northwest of Germany in the northwestern corner of the district of Cloppenburg, which is marked in gray on the map on the left. The four villages of the Saterland are shown in the map on the right.*

As reported by the literature, Saterland Frisian has a complete set of close short vowels: /i y u/ (Kramer 1968, Kramer 1982, Fort 2015). Together with the half-close short vowels /ɪ ʏ ʊ/ and the long close vowels /iː yː uː/ they constitute series of phonemes that differ by length and/or tenseness. The aim of this paper is to identify acoustic cues which distinguish the words within the triplets /ɪ i iː/ and /ʊ u uː/.

The acoustic cues, which add to the distinction of vowels in Saterland Frisian triplets have not yet been fully studied. Potential acoustic cues, which distinguish the vowels in a triplet, may be vowel duration and spectral features (F1, F2). Siebs' (1889) distinction between *Stoßton* and *Schleifton* in Saterland Frisian suggests that f0 might play an additional role even if Tröster-Mutz (1997, 2002) did not find evidence for tone accent differences in present-day Saterland Frisian (for further discussion see Peters 2008).

Several studies show that vowels with a stronger f0 dynamic are perceived as being longer. Among others, Lehiste (1976) 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 (2010) 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. This effect is likely to be language-specific (Cumming 2011, Lehnert-LeHouillier 2010).

In addition to normal speech, we will also consider clear speech. Rogers et al. (2010) define clear speech as "a speaking style that is used by talkers when they know they may have difficulty being understood, such as in noisy environments and when talking with hearing impaired persons." Several studies suggest that contrasts between vowels in duration and spectral properties may increase in clear speech.

Uchanski (1988, 1992) found that the duration contrast between tense/long and lax/short vowels was enhanced in English clear speech by lengthening the tense/long vowels to a greater extent than the lax/short vowels. This finding suggests that the nonuniform increase in segment durations for clear speech reflects the temporal structure of the language at the segmental level.

Bradlow (2002) examined the acoustic-phonetic modifications that characterize clear speech considering the high vowels /i/ und /u/. She especially focused on the extent of CV coarticulation and vowel space expansion as a function of vowel inventory size. Subjects were English monolinguals and Spanish-English bilinguals. Bradlow found a consistent increase of duration for both speaker groups and both languages. In clear speech the effects for CV coarticulation were maintained, and the vowel space was expanded both for English and Spanish. However, no correlation with the vowel inventory size was found.

Similarly, Smiljanić & Bradlow (2005) elicited conversational and clear speech from five native speakers of Croatian and five native speakers of English. The authors found that the acoustic-phonetic features of the conversational-to-clear speech transformation revealed cross-language similarities in clear speech production strategies. In both languages speakers exhibited a decrease in speaking rate and an increase in pitch range, as well as an expansion of the vowel space.

In summary, the studies show that clear speech may increase duration and durational contrast, pitch, pitch range, and vowel space size, which is confirmed by many other studies as well (cf. Picheny et al. 1986, 1989; Krause and Braida 2004; Bradlow et al. 2003; Liu et al. 2004; Moon and Lindblom 1994; Ferguson and Kewley-Port 2002; Johnson et al. 1993). A listener-directed clear speech task maximizes the discrimination between words and appears to be a promising approach for finding segmental and prosodic cues distinguishing the sounds within a triplet which may remain undetected in a normal speech experiment.

The present paper is an extended version of Heeringa et al. (2014). We investigate, which acoustic cues distinguish the sounds within triplets containing /ɪ i iː/ and /ʊ u uː/. The triple /ɪ i iː/ is also found in West Frisian (Fort 1980, p. 62, Tiersma 1999, p. 9, Visser 2003, p. 132). The distinction between high vowels /i iː/, /y yː/ and /u uː/ has been studied by de Graaf (1985) for West Frisian. His results were obtained on the basis of five subjects between 60 and 70 years old and five subjects between 20 and 30 years old and suggest that the distinction between short and long tense vowels decreases. The average vowel duration ratio for the older group was 2.4, whereas the younger group showed a ratio of 1.9. De Graaf also found that all speakers pronounced the short vowels in a more central position than the long vowels. For the younger generation the distances between the short and long tense vowels in the F1/F2 plane are systematically smaller than the corresponding distances for the older generation. This may give rise to the question whether short lax and tense high vowels are still distinguished.

In Saterland Frisian only for the trifold distinctions /ɪ i iː/ and /ʊ u uː/ minimal triplets are available and still known (cf. Fort (1980), p. 62), even if only by a restricted number of Saterland Frisian speakers.

In order to find the acoustic cues that distinguish the sounds within Saterland Frisian triplets we obtained normal speech by conducting a traditional reading task, and clear speech by carrying out a listener-directed task for each of the two triplets. Additionally, we study the effect of clear speech on the acoustic variables compared to normal speech. In particular, we address the following research questions:

1. Which acoustic cues distinguish the sounds within triplets containing /ɪ i iː/ and /ʊ u uː/ in normal speech?
2. Which acoustic cues distinguish the same triplets in clear speech? What is the effect of clear speech on the acoustic variables compared to normal speech?

3. Do native speakers of Saterland Frisian perceive the phonological distinctions within a triplet in clear speech?

In Section 2 we describe the normal speech experiment. Section 3 presents the results of the clear speech experiment. Results of the perception experiment are presented in Section 4. In Section 5 we end with conclusions and discussion.

2. Experiment 1: Normal speech

2.1. Material

We examined the /ɪ i iː/ and /ʊ u uː/ triples because for these trifold distinctions minimal triplets are available. For the closed front vowels /ɪ i iː/ we used the triplet *Smitte* [smɪtə] 'forge', *smiete* [smitə] 'to throw', and *Smíete* [smiːtə] 'throws' (pl.), which was preferred over the alternative triplet *linnen* 'made out of linen, linen', *Lienen* 'leashes' (pl.), and *Líenen* 'railings, backrests' due to greater familiarity of the former triplet and better segmentability. The closed back vowels /ʊ u uː/ were elicited by the only available and well-known triplet *ful* [fʊl] 'full', *fuul* [ful] 'rotten', and *fúul* [fuːl] '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 in 2012 with four female native speakers, aged 78, 66, 67, and 65 years, henceforth referred to as subject 1, subject 2, subject 3, and subject 4 respectively. The four speakers are early trilingual speakers of Saterland Frisian, Low German, and High German. They 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 (Fort 1980). Given the small number of speakers, this study should be considered as explorative.

Saterland Frisian words were presented in written form to each of the four native speakers on a computer screen one word a time. We used twelve different words: six triplet words (*ful*, *fuul*, *fúul*, *Smitte*, *smiete*, *Smíete*) and six filler words (*Pot* 'pot', *Paad* 'path', *Kat* 'cat', *leet* 'late', *Täk* 'roof', *Poot* 'paw').

A session consisted of four blocks in which each of the 12 words was presented four times. Within each block the words were presented in controlled randomized order, 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. Subject 1 performed the test twice, therefore we obtained 32 samples per subject per triplet word for this informant. The second test was done in order to obtain a larger set of samples with a f0 peak, i.e. of samples not produced with a downstepped high tone. The number of word samples is given per triplet and per subject in Table 1. When looking for cues concerning the f0 dynamics only samples with a clear f0 peak were used for the analysis. The number of word samples which satisfy this condition is given separately in Table 1. Note that for subjects 3 and 4 we did not obtain samples of *Smitte*, *smiete*, and *Smíete*, and no samples with f0 peak. Therefore, when results for the /ɪ i iː/ triplet are shown or when results concerning the f0 peak are shown, no results are given for subjects 3 and 4.

Table 1. *Number of word samples per triplet word and per subject obtained by the normal speech experiment.*

	all samples				samples with f0 peak			
	subject 1	subject 2	subject 3	subject 4	subject 1	subject 2	subject 3	subject 4
Smitte	32	16			25	12		
smiete	32	16			26	16		
Smíete	31	16			23	16		
ful	32	16	16	16	6	11		
fuul	32	16	16	16	10	13		
fúul	32	16	16	16	11	14		

2.3. Acoustic variables

Segmental and prosodic variables were measured with PRAAT (Boersma & Weenink, 1992-2015). For each word belonging to the /ɪ i iː/ triplet we measured the duration of each of the segments /s/, /m/, V, /t/, and 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 vowel center. When looking for acoustic cues that distinguish triplet words, we took the perception of the speakers as the starting point. To this end, Hertz values were converted to the auditory Bark scale using Traunmüller's (1990) formula in all of the analyses in this paper.

To analyze formant dynamics, the amount of vowel inherent spectral change was assessed as the trajectory length (TL), which is obtained on the basis of spectral changes in the vowel interval between the 20% point and the 50% point and between the 50% point and the 80% point of the vowel's duration. These spectral changes are calculated as Euclidean distances in Hertz between the measurement points in the F1-F2 plane following Fox & Jacewicz 2009 (cf. Jin & Liu 2013):

$$VSL_{50-20} = \sqrt{(F1_{50} - F1_{20})^2 + (F2_{50} - F2_{20})^2}$$
$$VSL_{80-50} = \sqrt{(F1_{80} - F1_{50})^2 + (F2_{80} - F2_{50})^2} \quad (1)$$

Subsequently TL is measured as the sum of the two vowel sections lengths (VSL50-20, VSL80-50).

$$TL = VSL_{50-20} + VSL_{80-50} \quad (2)$$

To account for dynamic changes in unnormalized time, the spectral rate of change (TL roc) between the 20% and 80% point, i.e. within the central 60%, was calculated as proposed in Fox & Jacewicz 2009 (cf. Mayr & Davies 2011):

$$TL\ roc = \frac{TL}{0.60 \times V_{dur}} \quad (3)$$

We measured f0 variation in the interval from the beginning of /m/ to the end of the vowel. Where a clear f0 peak was present within this interval, we measured the rise size and the fall size in semitones (see Figure 2). Relative f0 excursion in semitones per millisecond was used as a measure of f0 dynamics. We calculated f0 excursion by dividing 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) by the duration of the interval.

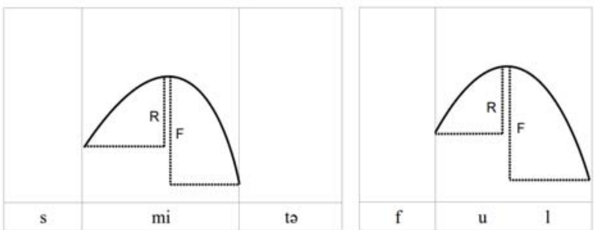


Figure 2. *Pitch contour of the triplet word smiete (left) and fuul (right). Vertical lines with R show the f0 rise size and vertical lines with F show the f0 fall size.*

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

2.4. Statistical processing

In order to measure average differences between triplet words, we used the robust *R* function pbtrmp, which does not require particular sample distributions and can deal with small sample sizes (cf. Table 1). The function pbtrmp was developed by Rand Wilcox (tmcppb uses a strongly related methodology and is discussed in Wilcox 2012). This function performs multiple comparisons for independent groups based on trimmed means and using a percentile bootstrap method. The amount of trimming is 20%. The estimate of a 20% trimmed mean is based on the 20% Winsorized sample variance. In a Winsorized sample outliers are replaced with observed values, rather than discarded. Each of the first k smallest values are replaced by the (k+1)-th smallest value, and the k largest values are replaced by the (k+1)-th largest value. The Winsorized variance of a sample is the variance of the winsorized set of values (Hastings, Mosteller, Tukey & Winsor 1947).

The acoustic 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 one of the subjects for a particular variable or a significant difference in opposite direction, no significance is reported.

In the tables below, statistical significant differences are measured at the 0.05 level. The direction of the differences is indicated by '<' (measurements for the first triplet word are smaller than for the second triplet word) and '>' (measurements for the first triplet word are larger than for the second triplet word).

In several analyses we determined how well (combinations of) acoustic variables predict the distinctive triplet words. For this purpose, we use linear discriminant analysis as implemented by the R function lda in the package MASS (Venables & Ripley 2002).

2.5 Cues distinguishing triplet words

Duration values for the /ɪ i iː/ triplet are shown in Figure 3 for subject 1 and subject 2. Vowel plots are found in Figure 4 and f0 contour plots in Figure 5.

/ɪ i iː/ triplet /ʊ u uː/ triplet

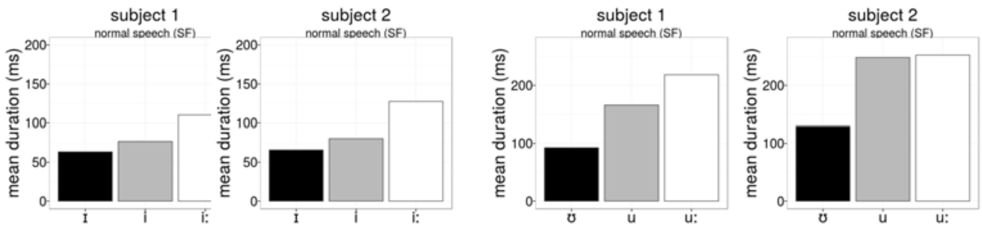


***Figure 3**. Duration values for normal speech of the /ɪ i iː/ triplet (top) and the /ʊ u uː/ triplet (bottom).*

/ɪ i iː/ triplet /ʊ u uː/ triplet

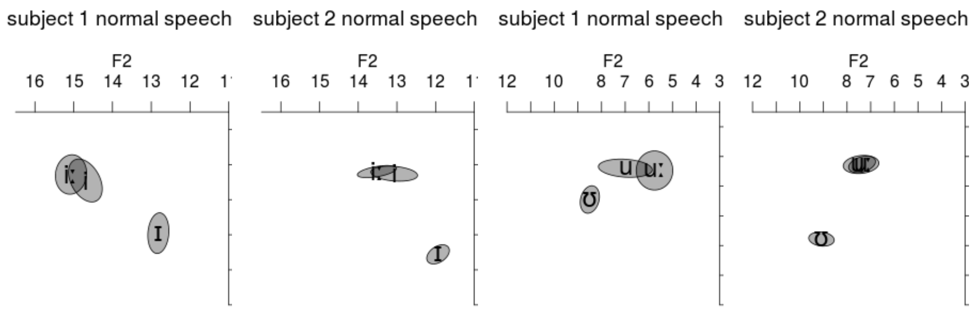


***Figure 4.** Vowel plots show the mean formant values of the triplet sounds of the /ɪ i iː/ triplet and the /ʊ u uː/ triplet for normal speech. Ellipses represent confidence intervals of 68.27% (equivalent to plus-or-minus 1 sample standard deviation). As expected from the depiction of the Saterland Frisian vowel system in (cf. Fort 1980, p.60) the short vowels /ɪ ʊ/ are produced more centralized than /iː i/ and /uː u/.*

/ɪ i iː/ triplet /ʊ u uː/ triplet

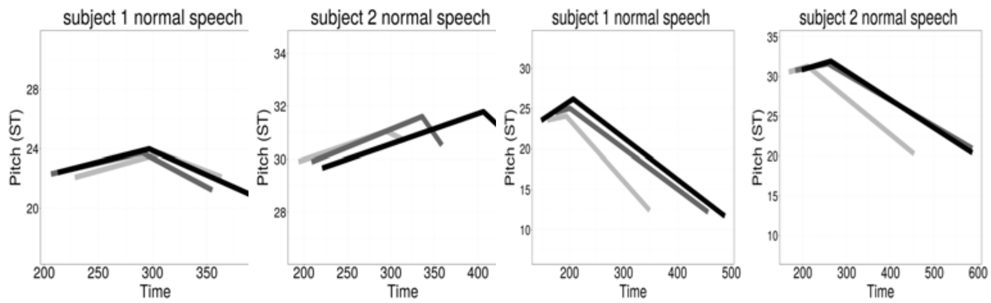


Figure 5. *f0 contours for the /ɪ i iː/ triplet and the /ʊ u uː/ triplet in normal speech. Light grey lines represent short lax vowels, dark gray lines represent short tense vowels, and black lines represent long tense vowels.*

Consensus results of the two subjects are presented in Table 2a. We find that short lax and tense vowels are distinguished by vowel duration, F1, F2, and f0 fall size. Short and long tense vowels are distinguished by vowel duration and F2.

Table 2. *Consensus results of subjects 1 and 2 for the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b), and of subjects 1, 2, 3, and 4 for the /ʊ u uː/ triplet (c).*

(a)	1-2	1-3	2-3	(b)	1-2	1-3	2-3	(c)	1-2	1-3	2-3
/s/ duration				/f/ duration				/f/ duration			
/m/ duration											
V duration	<	<	<	V duration	<	<		V duration			
t1 duration				/l/ duration				/l/ duration			
t2 duration											
/ə/ duration											
F1	>	>		F1	>	>		F1	>	>	
F2	<	<	<	F2	>	>		F2	>	>	
TL				TL				TL			
TL roc		>		TL roc	>			TL roc			
f0 rise size				f0 rise size							
f0 fall size	<	<		f0 fall size							
f0 dynamics				f0 dynamics							

In (a) 1=Smitte, 2=smiete, 3=Smíete, and in (b) and (c) 1=ful, 2=fuul, 3=fúul.

Duration values for the /ʊ u uː/ triplet are shown in Figure 3 for subject 1 and subject 2. Vowel plots are found in Figure 4 and f0 contour plots in Figure 5.

Consensus results of the two subjects are presented in Table 2b. Short lax and tense vowels are distinguished by vowel duration, F1 and F2. Short lax and tense vowels are also distinguished by TL roc. The triplet words fuul and fúul are not distinguished by any variable, i.e. short tense and long tense vowels are not distinguished in normal speech.

The /ʊ u uː/ triplet was also studied for all four subjects. Since the number of samples with a clear f0 peak was not sufficient for subject 3 and subject 4, we studied the triplet without considering variables regarding f0. Duration values for the /ʊ u uː/ triplet are shown in Figure 6 for all four subjects. Vowel plots are found in Figure 7.

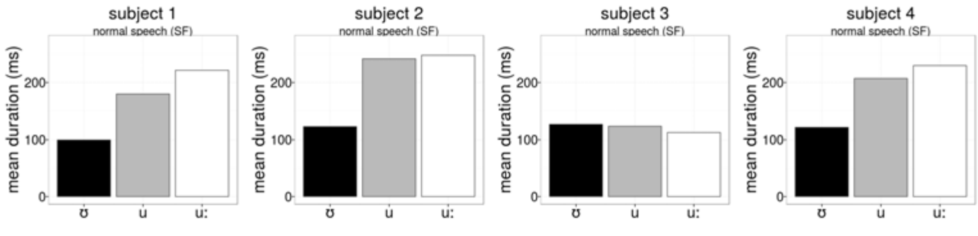


Figure 6. *Duration values for normal speech for the /ʊ u uː/ triplet.*

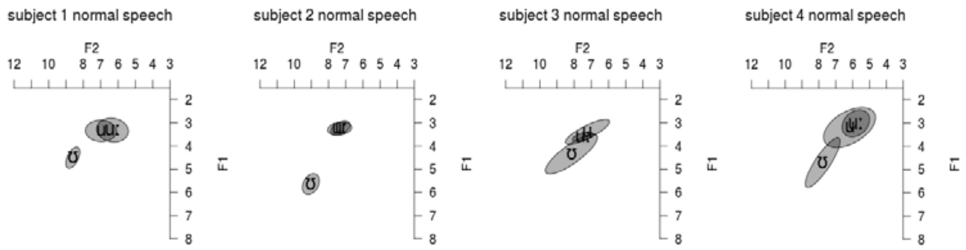


Figure 7. *Vowel plots show the mean formant values of the triplet sounds for normal speech using Saterland Frisian stimuli. Ellipses represent confidence intervals of 68.27% (equivalent to plus-or-minus 1 sample standard deviation). As expected from the depiction of the Saterland Frisian vowel system in (cf. Fort 1980, p.60) the short vowels /ɪ ʊ/ are produced more centralized than /iː i/ and /uː u/.*

Consensus results of the four subjects are presented in Table 2c. We find that short lax vowels are distinguished from short and long tense vowels by F1 and F2. The triplet words fuul and fúul are not distinguished by any variable.

For all three data sets we find that short lax and tense vowels are distinguished by F1 and F2. In the /ɪ i iː/ triplet short and long tense vowels are distinguished by vowel duration and F2, but in the /ʊ u uː/ triplet this distinction is not found for any variable.

2.6 Cues predicting triplets

2.6.1 Prediction of lax/tense and short/long distinction

Using linear discriminant analysis, we tried to find the variables that best predict the distinction between short lax and tense vowels, and between short and long tense vowels. Table 3a shows the acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ɪ i iː/ triplet. When we consider the lax/tense distinction we find highest percentages for F2, and second best for F1. The short/long distinction is best predicted by vowel duration.

Table 3. *Acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b) with linear discriminant analysis. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

(a)	lax versus tense		short versus long	
	subj 1	subj 2	subj 1	subj 2
/s/ duration	57	61	47	50
/m/ duration	51	61	61	72
V duration	73	82	**88**	**84**
t1 duration	73	68	55	53
t2 duration	67	64	47	50
/ə/ duration	45	57	65	56
F1	94	**100**	63	56
F2	**100**	**100**	76	75
TL	53	61	65	56
TL roc	55	64	76	50
f0 rise size	57	68	59	66
f0 fall size	78	82	65	63
f0 dynamics	55	61	53	69

(b)	lax versus tense		short versus long	
	subj 1	subj 2	subj 1	subj 2
/f/ duration	63	54	67	63
V duration	81	96	76	52
/l/ duration	56	54	52	63
F1	**100**	**100**	48	52
F2	94	96	**86**	56
TL	69	54	67	67
TL roc	63	88	57	63
f0 rise size	63	54	62	59
f0 fall size	63	54	62	48
f0 dynamics	63	79	52	63

For the /ʊ u uː/ triplet results are shown in Table 3b. When we consider the lax/tense distinction, we find the highest percentages for F1, and second highest percentages for F2. Regarding the short/long distinction we find that there is no consensus for the two subjects. For subject 1 F2 has the highest percentage, and for subject 2 TL has the highest percentage. In Table 2b in section 2.5, we found that there is likewise no consensus about which variable distinguishes short and long tense vowels.

The results on the basis of four speakers in Table 4 are similar. The lax/tense distinction is best predicted by F1 and second best by F2, but there is no consensus among the four speakers about which variable predicts the short/long distinction best. In Table 2c we find that there is also no consensus about which variable distinguishes short and long tense vowels.

Table 4. *Acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

	lax versus tense				short versus long			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
/f/ duration	61	56	53	63	61	**66**	59	66
V duration	77	97	59	88	64	53	56	53
/l/ duration	63	50	69	66	58	53	66	**72**
F1	**98**	**100**	**88**	**91**	45	50	**69**	59
F2	95	97	72	88	**70**	59	59	56
TL	52	53	44	63	55	59	53	41
TL roc	72	75	56	66	59	**66**	47	56

The results suggest that the distinction between short lax and tense vowels is predicted by spectral variables, F1 and/or F2 for all subjects and both triplets. Within the /ɪ i iː/ triplet tense short and long tense vowels are predicted by vowel duration. For the /ʊ u uː/ triplet for each subject another variable is found that predicts the distinction between short and long tense vowels best. Note that for none of the subjects vowel duration was found to be the best predictor.

2.6.2 Prediction of triplet words by one variable

In the previous section we looked for variables that predict the lax/tense distinction, and for variables that predict the short/long distinction. In this section we study how well individual variables predict all three vowels within a triplet. Results for the /ɪ i iː/ triplet are shown in Table 5a. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentages are found for F2.

Table 5*. Acoustic variables predicting the words within the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b). The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

(a)	subject 1	subject 2
V duration	73	73
F1	73	68
F2	**84**	**82**
TL	42	41
f0 rise size	39	50
f0 fall size	62	64

(b)	subject 1	subject 2
V duration	67	61
F1	59	**66**
F2	**85**	**66**
TL	56	47
f0 rise size	48	45
f0 fall size	48	37

For the /ʊ u uː/ triplet results are shown in Table 5b. For subject 1 F2 is the best predictor, for subject 2 both F1 and F2 have the highest percentage.

Table 6 shows the results for the /ʊ u uː/ triplet on the basis of four speakers and all available samples per speaker. As in Table 5b, for subjects 1 and 2 we find that F2 is the best predictor but for subjects 3 and 4 the best predictor is F1.

Table 6*. Acoustic variables predicting the words within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

	subject 1	subject 2	subject 3	subject 4
V duration	60	67	44	60
F1	63	67	**71**	**67**
F2	77	**71**	60	63
TL	38	40	40	33

For both triplets we find that spectral features best predict triplet words. We find that the vowels within the /ɪ i iː/ triplet are predicted relatively well by just one predictor, namely F2. For the /ʊ u uː/ triplet percentages are lower on average and there is no consensus among the subjects whether F1 or F2 is the best predictor.

2.6.3 Prediction of triplet words by combinations of variables

In Section 2.6.2 we found that vowels within the /ɪ i iː/ triplet are predicted relatively well by just one predictor. However, using just one predictor for the /ʊ u uː/ triplet was less successful. In Section 2.6.1 we found that the distinction between short lax and tense vowels is predicted by spectral variables, and within the /ɪ i iː/ triplet short and long tense vowels are predicted by vowel duration. This might indicate that at least two different variables are required for an optimal prediction of the vowels within a triplet.

Table 7a shows combinations of two acoustic variables predicting the words within the /ɪ i iː/ triplet with linear discriminant analysis. The highest percentages are found for vowel duration + F2, followed by vowel duration + F1. This agrees with the results we found in Section 2.6.1, where spectral properties were found to best predict the lax/tense distinction and vowel duration was the best predictor of the short/long distinction.

Table 7. *Combinations of two acoustic variables predicting the words within the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b). The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

(a)		subj 1	subj 2	(b)		subj 1	subj 2
V duration	F1	88	**91**	V duration	F1	78	68
V duration	F2	**92**	**91**	V duration	F2	85	66
V duration	TL	69	75	V duration	TL	74	71
V duration	f0 rise size	72	68	V duration	f0 rise size	70	68
V duration	f0 fall size	76	80	V duration	f0 fall size	74	68
F1	F2	82	82	F1	F2	**89**	68
F1	TL	72	70	F1	TL	78	**76**
F1	f0 rise size	70	73	F1	f0 rise size	78	66
F1	f0 fall size	76	80	F1	f0 fall size	74	68 →

F2	TL	86	84	F2	TL	85	74
F2	f0 rise size	85	84	F2	f0 rise size	85	68
F2	f0 fall size	86	80	F2	f0 fall size	85	68
TL	f0 rise size	46	45	TL	f0 rise size	63	58
TL	f0 fall size	64	64	TL	f0 fall size	67	42
f0 rise size	f0 fall size	61	68	f0 rise size	f0 fall size	56	37

Results for the /ʊ u uː/ triplet are given in Table 7b. We could not find a consensus among the subjects. The best combination for subject 1 is F1 + F2, and for subject 2 F1 + TL. This agrees with the results in Section 2.6.1, where the lax/tense distinction was predicted best by F1, and regarding the short/long distinction F2 had the highest percentage for subject 1, and TL had the highest percentage for subject 2. Since no consensus could be found for the variable predicting the short/long distinction best (see Table 3b), no consensus combination can be found that predicts all three triplet words best.

Results obtained on the basis of four speakers are shown in Table 8. We could not find consensus for the best combination of variables across the four speakers. This can be explained from the fact that no consensus could be found for the short/long distinction (see Table 4).

Table 8. *Combinations of two acoustic variables predicting the words within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

		subject 1	subject 2	subject 3	subject 4
V duration		60	67	44	60
V duration	F1	77	67	75	63
V duration	F2	72	71	65	63
V duration	TL	59	73	46	**69**
F1		63	67	71	67
F1	F2	**78**	65	77	69
F1	TL	70	73	69	60
F2		77	71	60	63
F2	TL	72	75	58	60
TL		38	40	40	33

To summarize we find consensus for the /ɪ i iː/ triplet, where triplet words are predicted best by a combination of F2 and vowel duration, but for the /ʊ u uː/ triplet a consensus combination cannot be found, likely because short tense and long tense vowels are not distinguished in normal speech (see Tables 2b and 2c). The length distinction may be neutralized in some contexts.

When comparing the results on the basis of two variables with the results on the basis of one variable (Section 2.6.2), we find the biggest improvement for the /ɪ i iː/ triplet: 82-84% versus 91-92%. Less improvement is found for the /ʊ u uː/ triplet when considering two speakers: 66-85% versus 76-89%. Hardly any improvement is found for the /ʊ u uː/ triplet when considering four speakers: 67-77% versus 69-78%.

3. Experiment 2: Clear speech

3.1 Procedure

Saterland Frisian words were presented in written form on a computer screen to the same four speakers who participated in experiment 1. 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 highlighted by blue colour (the other words were black). The three words of a triplet were located on the screen so that they are imaginary vertices of a triangle. Each triangle was rotated over an arbitrary angle. Screenshots of the clear speech condition used to obtain samples of the triplet words *ful*, *fuul*, and *fúul* are exemplified in Figure 8.

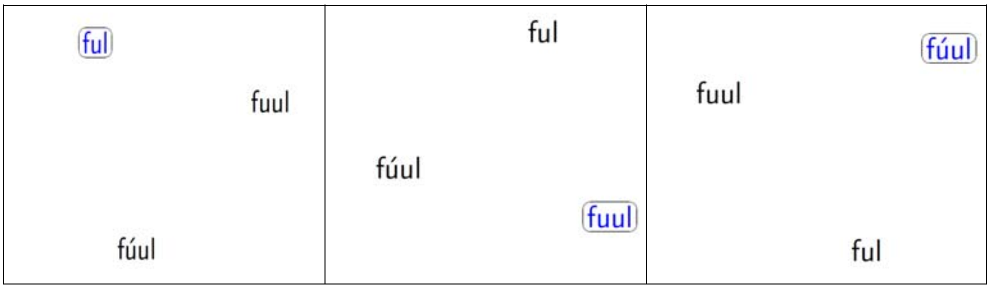


Figure 8. *Sequence of three screenshots in a clear speech experiment.*

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. A screen separated the reader and the listener during this experiment.

Just as for normal speech we obtained 16 samples per subject per triplet word. The number of word samples is given per triplet and per subject in Table 9. When we look 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 separately in Table 9. Note that for subjects 3 and 4 we did not obtain samples of Smitte, smiete, and Smíete, and no samples with an f0 peak. Therefore, when results for the /ɪ i iː/ triplet are shown or when results concerning the f0 peak are shown, no results are given for subjects 3 and 4.

Table 9. *Number of word samples per triplet word and per subject obtained by the clear speech experiment with Saterland Frisian stimuli. The right half of the table shows the number of samples which have a clear f0 peak.*

	all samples				samples with f0 peak			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
Smitte	16	16			14	5		
smiete	16	16			16	14		
Smíete	16	16			16	14		
ful	16	16	14	11	5	4		
fuul	16	16	13	13	4	10		
fúul	16	16	14	12	13	10		

We conducted a second clear speech experiment, in which we tried to avoid a possible influence of the generally unfamiliar Saterland Frisian spelling on the subject. When a subject reads *fuul* and *fúul* in succession, it is likely that s/he will pronounce *fúul* with longer vowel duration than fuul, even if the subject is not used to pronounce the two words distinctly. In this second experiment High German (HG) equivalents were presented instead of the Saterland Frisian (SF) words. The subjects pronounced the Saterland Frisian translations of the words, i.e. they translated and pronounced the words at the same time.

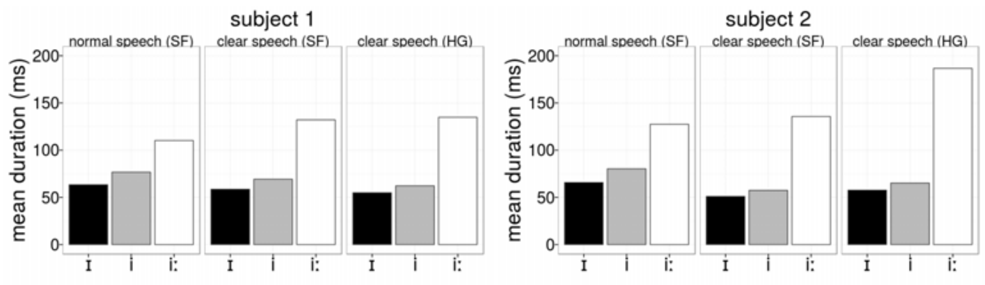
We obtained 16 samples per subject per triplet word. The number of word samples is given per triplet and per subject in Table 2b. When we look for cues concerning the f0 dynamics, only samples with a clear f0 peak can be used for the analysis. The number of word samples, which satisfy this condition, is given separately in Table 10. Samples with f0 were obtained only for subjects 1 and 2 for *Smitte*, *smiete*, and *Smíete*.

Table 10. *Number of word samples per triplet word and per subject obtained by the clear speech experiment with High German stimuli.*

	all samples				samples with f0 peak			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
Smitte	16	16			16	8		
smiete	16	16			15	11		
Smíete	16	15			16	9		
ful	16	15	18	18				
fuul	16	16	18	18				
fúul	16	16	18	18				

3.2 Cues distinguishing triplet words

Duration values for the /ɪ i iː/ triplet as obtained by the clear speech experiment using Saterland Frisian stimuli are shown in Figure 9 for subject 1 and subject 2. Vowel plots are found in Figure 10 and f0 contour plots in Figure 11.



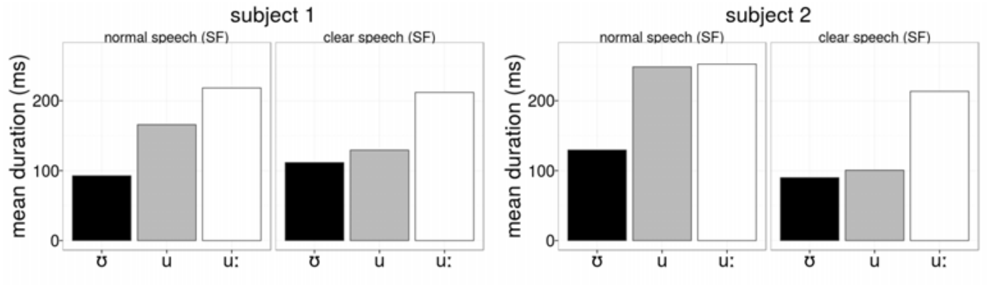
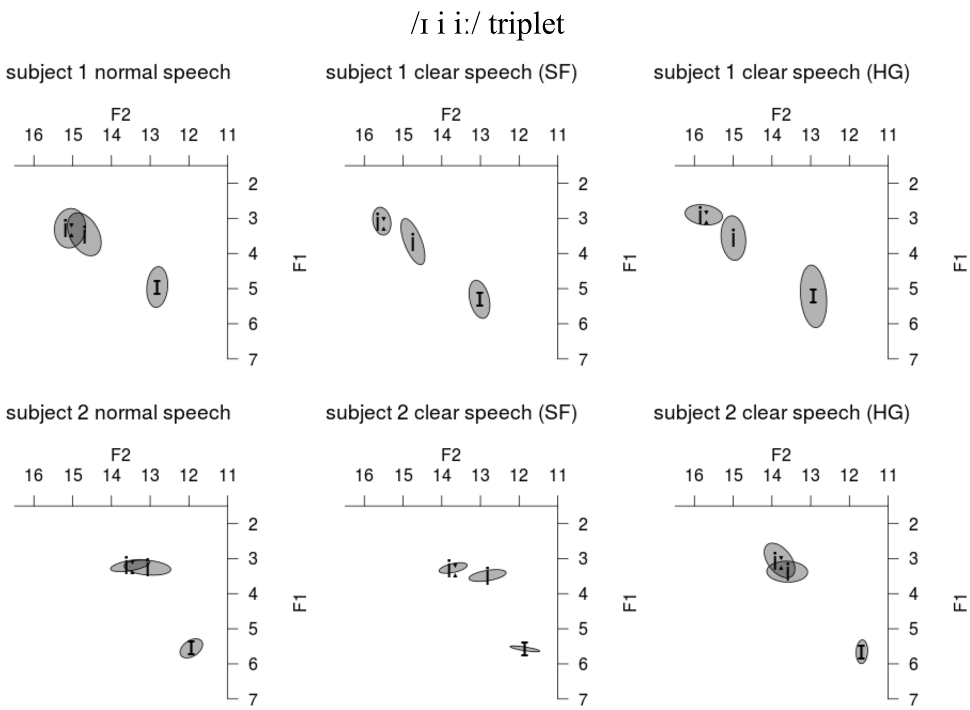


Figure 9. *Duration values for normal speech and clear speech. On top the /ɪ i iː/ triplet and at the bottom the /ʊ u uː/ triplet.*



/ʊ u uː/ triplet

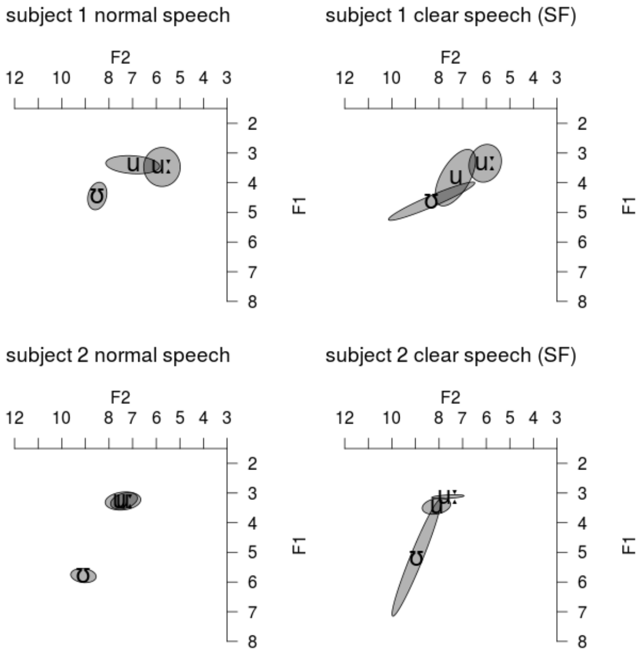
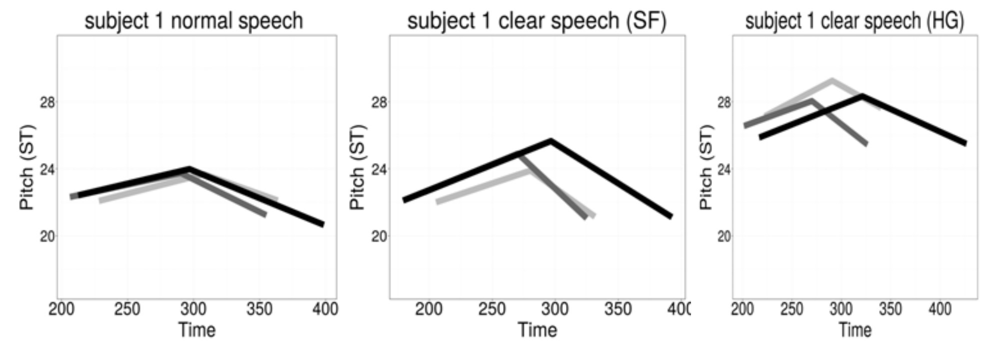
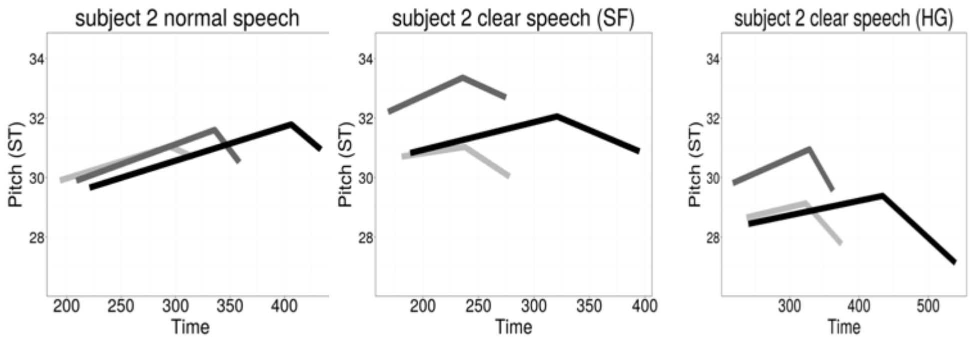


Figure 10. *Vowel plots show the mean formant values of the triplet sounds of the /ɪ i iː/ triplet (top) and the /ʊ u uː/ triplet (bottom) for normal speech (left), clear speech using Saterland Frisian stimuli (middle), and clear speech using High German stimuli (right). Ellipses represent confidence intervals of 68.27% (equivalent to plus-or-minus 1 sample standard deviation).*

/ɪ i iː/ triplet





/ʊ u uː/ triplet

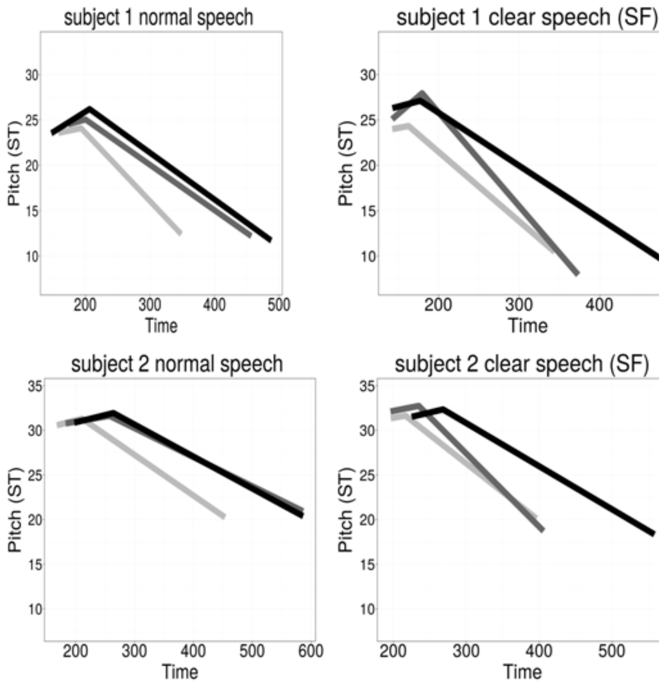


Figure 11. *f0 contours for the /ɪ i iː/ triplet (top) and the /ʊ u uː/ triplet (bottom) in normal speech (left), clear speech using Saterland Frisian stimuli (middle), and clear speech using High German stimuli (right). Lighter lines represent short lax vowels, darker gray lines represent short tense vowels, and black lines represent long tense vowels.*

Consensus results of the two subjects on the basis of normal speech and clear speech are presented in Table 11a. For both normal speech and clear speech all three triplet words are distinguished from each other by vowel duration. The distinctions in normal speech found for F1 and TLroc are also found in the two types of clear speech. When comparing clear speech to normal speech, we find that a larger number of variables distinguish short vowels and long vowels. In addition to vowel duration, F1 and F2, short and long tense vowels are distinguished by durations of /m/, t2, and /ə/, and by TLroc. Differences between clear speech elicited by Saterland Frisian stimuli and clear speech elicited by High German stimuli are small. When Saterland Frisian stimuli were used, short lax and long tense vowels are distinguished by f0 rise size. However, short lax and short and long tense vowels are distinguished by f0 fall size when High German stimuli were used. Using Saterland Frisian stimuli F2 distinguished short and long tense vowels and f0 dynamics distinguished short lax and tense vowels. When using High German stimuli this is no longer the case.

The results for f0 are different for each type of speech. But in both normal speech and clear speech (HG) short lax and long tense vowels are distinguished by f0 fall size, and in the two types of clear speech short tense and long tense vowels are distinguished by f0 dynamics.

Table 11. *Consensus results of subject 1 and subject 2 for the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b).*

(a)	normal speech			clear speech (SF)			clear speech (HG)		
	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3
/s/ duration									
/m/ duration					<	<		<	<
V duration	<	<	<	<	<	<	<	<	<
t1 duration									
t2 duration					<	<		<	<
/ə/ duration					<	<		<	<
F1	>	>		>	>	>	>	>	>
F2	<	<	<	<	<	<	<	<	
TL									
TL roc		>			>	>		>	>
f0 rise size					<				
f0 fall size	<	<						<	
f0 dynamics				<		>			>

(b)	normal speech			clear speech		
	1-2	1-3	2-3	1-2	1-3	2-3
/f/ duration						
V duration	<	<			<	<
/l/ duration						
F1	>	>		>	>	
F2	>	>			>	
TL						
TL roc	>				>	
f0 rise size						
f0 fall size						
f0 dynamics						>

In (a) 1=Smitte, 2=smiete, 3=Smíete, and in (b) 1=ful, 2=fuul, 3=fúul.

Duration values for the /ʊ u uː/ triplet as obtained by the clear speech experiment using Saterland Frisian stimuli are shown in Figure 3 for subject 1 and subject 2. Vowel plots are found in Figure 12 and f0 contour plots in Figure 13.

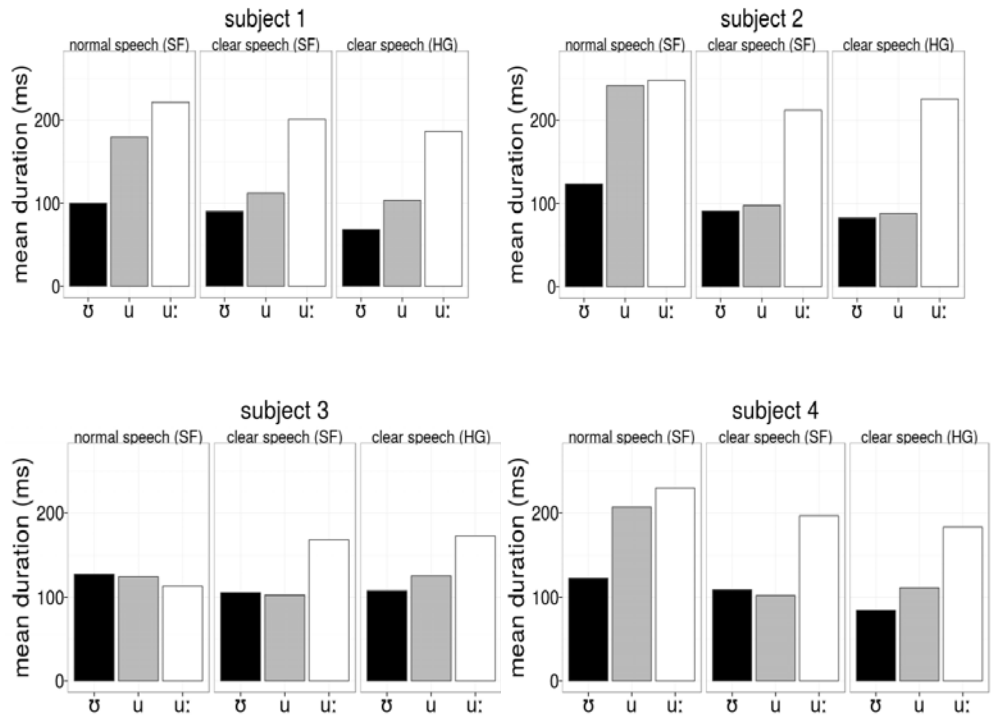


Figure 12. *Duration values for normal speech and clear speech for the /ʊ u uː/ triplet.*

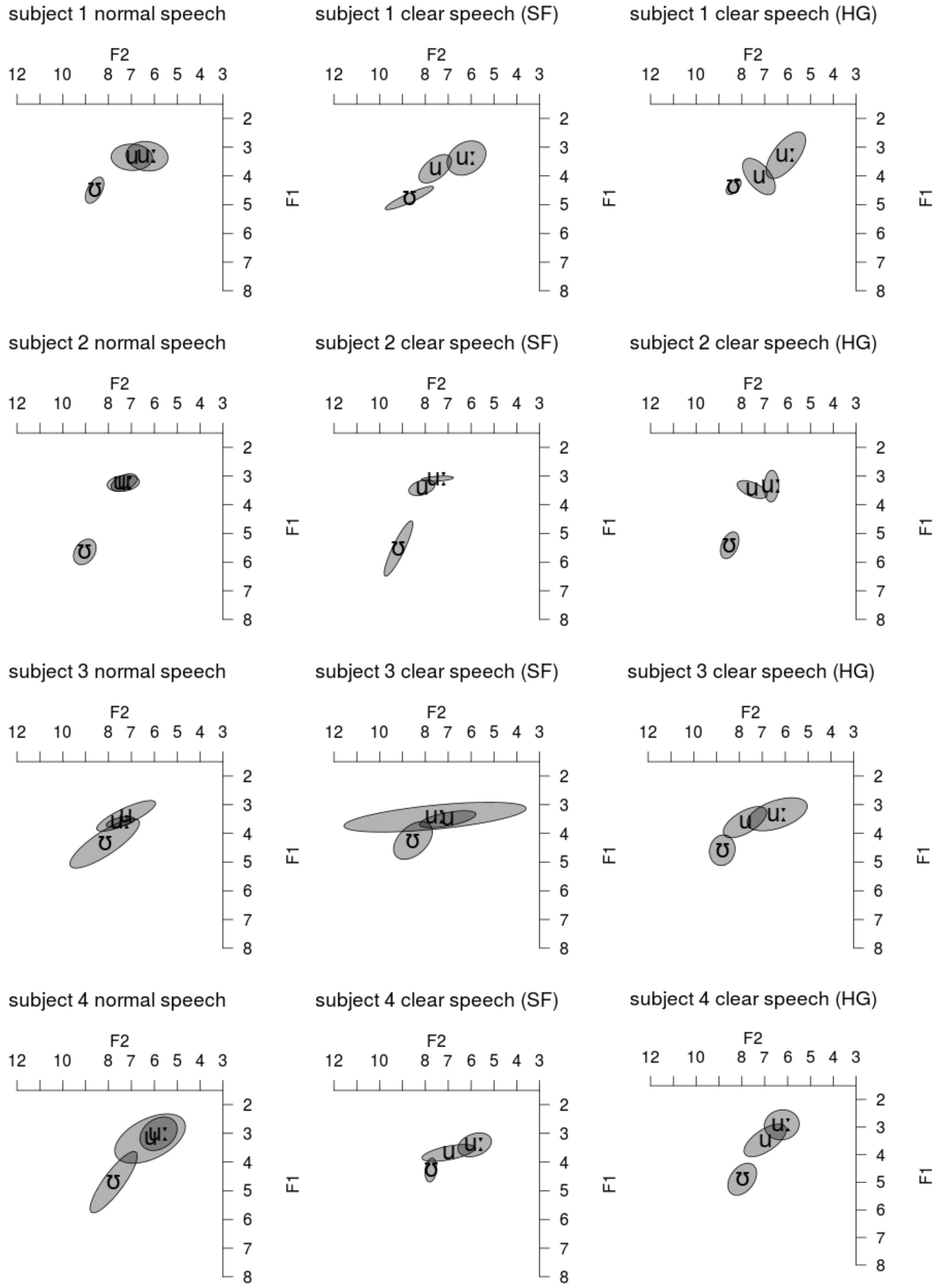


Figure 13. *Vowel plots show the mean formant values of the triplet sounds for normal speech (left), clear speech using Saterland Frisian stimuli (middle), and →*

clear speech using High German stimuli (right). Ellipses represent confidence intervals of 68.27% (equivalent to plus-or-minus 1 sample standard deviation).

Consensus results for normal speech and clear speech of the two subjects are presented in Table 11b. Normal speech and clear speech agree concerning F1, which distinguishes short lax vowels from short tense vowels. When comparing clear speech to normal speech, we find that the number of variables that distinguish short lax and tense vowels has decreased, in clear speech they are distinguished only by F1. In normal speech short and long tense vowels were not distinguished. In clear speech they are distinguished by two variables: vowel duration and f0 dynamics.

Duration values for the /ʊ u uː/ triplet are shown in Figure 11 for all four subjects. Vowel plots are found in Figure 12. Consensus results of the four subjects for the normal speech and the two clear speech experiments are presented in Table 12. In all three types of speech, short lax and tense vowels are distinguished by F1 and F2. When comparing clear speech to normal speech, we find that short and long tense vowels are also distinguished. When using Saterland Frisian stimuli, this distinction is represented by just one variable – V duration – but when using High German stimuli, three extra variables represent this distinction: /l/ duration, F2, and TL roc.

Table 12. *Consensus results of subject 1, subject 2, subject 3, and subject 4 for the /ʊ u uː/ triplet.*

	normal speech			clear speech (SF)			clear speech (HG)		
	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3
/f/ duration									
V duration					<	<		<	<
/l/ duration									<
F1	>	>		>	>		>	>	
F2	>	>		>	>		>	>	>
TL									
TL roc									>

1=ful, 2=fuul, 3=fúul

When considering the results of both triplets, we conclude that the effect of clear speech compared to normal speech is that the distinction between short and long tense vowels is revealed (/ʊ u uː/ triplet) or is revealed in a larger number of variables (/ɪ i iː/).

The effect of presenting High German stimuli instead of Saterland Frisian stimuli in the clear speech experiment seems to have a different effect on the two triplets. For the /ɪ i iː/ triplet some variables lost their significance for some distinctions but for the /ʊ u uː/ triplet the number of variables that distinguish short and long tense vowels increased from one to four. In order to try to explain the difference in effect on the triplets, we consider that translating the High German words into Saterland Frisian is harder than just reading the Saterland Frisian words. This is especially the case for the /ɪ i iː/ triplet, where the subjects had to translate the German words *werfen*, *Würfe*, and *Schmiede* into Saterland Frisian *Smitte*, *smiete*, and *Smíete* respectively. When translating the German words *voll*, *faul*, and *viel* into Saterland Frisian *ful*, *fuul* and *fúul* the subjects may have felt more confident, since all of the German words are cognates of the Saterland Frisian words, while the different pronunciation of the nucleus for each of the three triplet words might have stimulated them to pronounce the Saterland Frisian words even more clearly. Finally, we have to take into account that the results for the /ʊ u uː/ triplet are obtained on the basis of a larger number of speakers and samples than the results for the /ɪ i iː/, thus smaller effects reach statistical significance more easily.

3.3 Cues predicting triplets

In Section 2.6 we used linear discriminant analysis in order to find the variables that predict best the distinction between short lax and tense vowels, and between short and long tense vowels for normal speech. In this section we also consider clear speech, obtained on the basis of both Saterland Frisian stimuli and High German stimuli. In Section 3.3.1 we find the variables that best predict the distinction between short lax and tense vowels, and between short and long tense vowels. In Section 3.3.2 we find the variables, which best predict all distinctions within a triplet. In Section 3.3.3 we consider combinations of two variables and try to find the combination, which optimally predicts all three vowels within a triplet.

3.3.1 Prediction of the lax/tense and short/long distinction

Table 13 shows the acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ɪ i iː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated by the procedure. Both for normal speech and clear speech we find the highest percentages for the spectral features F1 and F2 when we consider the lax/tense distinction, and the highest percentages for duration when considering the short/long distinction. For subject 1 we also find high percentages for /ə/ duration and TL roc in the two types of clear speech.

For f0 we do not find any clear pattern. None of the three variables systematically performs better than the others, there is not a clear difference for the lax/tense and short/long distinction among the three types of speech, and the percentages are relatively low, varying between 48% and 82%.

Table 13. *Acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ɪ i iː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

	lax versus tense						short versus long					
	normal speech		clear speech (SF)		clear speech (HG)		normal speech		clear speech (SF)		clear speech (HG)	
	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2
/s/ duration	57	61	60	84	61	58	47	50	59	75	61	65
/m/ duration	51	61	57	74	48	58	61	72	81	79	84	85
V duration	73	82	77	84	61	79	**88**	**84**	**100**	**100**	**100**	**100**
t1 duration	73	68	87	74	55	58	55	53	56	82	61	75
t2 duration	67	64	77	74	61	58	47	50	69	75	77	90
/ə/ duration	45	57	60	74	58	58	65	56	94	86	90	80
F1	94	**100**	97	**100**	97	100	63	56	75	82	87	90
F2	**100**	**100**	**100**	95	**100**	**100**	76	75	**100**	96	87	65
TL	53	61	57	74	61	58	65	56	72	64	87	55
TL roc	55	64	53	74	52	47	76	50	91	82	94	70
f0 rise size	57	68	57	79	68	74	59	66	59	50	77	60
f0 fall size	78	82	70	68	77	58	65	63	59	61	48	70
f0 dynamics	55	61	80	68	48	68	53	69	66	71	74	90

For the /ʊ u uː/ triplet results are shown in Table 14. Considering the lax/tense distinction, we find the highest percentages for F1. For subject 1, the highest percentage (78%) is shared with f0 fall size. Regarding the short/long distinction, we find that there is no consensus for the two subjects in normal speech. However, for clear speech there is a clear consensus: short and long tense vowels are distinguished by vowel duration. For subject 1, the largest percentage (94%) is shared with F2 and f0 dynamics. The results give the impression that f0 distinguishes the triplet words better for subject 1 than for subject 2.

Table 14. *Acoustic variables predicting the lax/tense distinction of short vowels and the short/long distinction of tense vowels within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

	lax versus tense				short versus long			
	normal speech		clear speech (SF)		normal speech		clear speech (SF)	
	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2
/f/ duration	63	54	56	71	67	63	76	70
V duration	81	96	56	64	76	52	**94**	**100**
/l/ duration	56	54	67	71	52	63	76	55
F1	**100**	**100**	**78**	**93**	48	52	76	90
F2	94	96	67	**93**	**86**	56	**94**	65
TL	69	54	67	79	67	67	76	70
TL roc	63	88	44	79	57	63	88	60
f0 rise size	63	54	67	71	62	59	82	65
f0 fall size	63	54	**78**	86	62	48	82	55
f0 dynamics	63	79	67	79	52	63	**94**	80

The results on the basis of four speakers are given in Tables 15 and 16. In nearly all cases we find that the lax/tense distinction is best predicted by F1 and second best by F2. There is no consensus among the four speakers on which variable predicts the short/long distinction best. However, in clear speech there is a clear consensus: vowel duration predicts this distinction best.

Table 15. *Acoustic variables predicting the lax/tense distinction of short vowels within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

	normal speech				clear speech (SF)				clear speech (HG)			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
/f/ duration	61	56	53	63	69	59	56	75	59	61	42	53
V /duration	77	97	59	88	84	56	59	63	88	65	53	75
/l/ duration	63	50	69	66	66	47	67	**96**	44	68	61	86
F1	**98**	**100**	**88**	**91**	**94**	**97**	**93**	**96**	69	**100**	**92**	**100**
F2	95	97	72	88	**94**	91	85	**96**	**94**	90	83	86
TL	52	53	44	63	66	78	81	50	84	39	61	72
TL roc	72	75	56	66	53	78	81	54	72	48	53	64

Table 16. *Acoustic variables predicting the short/long distinction of tense vowels within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

	normal speech				clear speech (SF)				clear speech (HG)			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
/f/ duration	61	**66**	59	66	56	63	70	60	66	84	64	47
V duration	64	53	56	53	**94**	**100**	**89**	**100**	**100**	**100**	81	**89**
/l/ duration	58	53	66	**72**	75	56	59	60	84	91	72	75
F1	45	50	**69**	59	78	91	70	68	81	59	69	78
F2	**70**	59	59	56	**94**	75	44	88	88	94	**89**	78
TL	55	59	53	41	53	75	52	60	63	59	56	78
TL roc	59	**66**	47	56	78	59	48	60	72	69	64	86

We conclude that for both triplets the lax/tense distinction is best predicted by spectral features F1 and F2, and the short/long distinction is best predicted by vowel duration. The effect of clear speech compared to normal speech is found for the short/long distinction in the /ʊ u uː/ triplet. While there is no consensus about speakers for normal speech, in clear speech we do find a consensus: vowel duration has the highest percentage of correctly predicted short tense and long tense vowels. Considering the two types of clear speech, no consistent difference between the use of Saterland Frisian stimuli and High German stimuli is evident.

3.3.2 Prediction of triplet words by one variable

Table 17a shows the acoustic variables, which predict the complete set of vowels within the /ɪ i iː/ triplet. Regarding normal speech and clear speech (SF) there is a clear consensus: F2 is the best predictor. For clear speech (HG) the consensus is that triplet words are best predicted by spectral features but there is no consensus about which spectral features are the best predictors. For subject 1 this would either be F2 or TL, for subject 2 this would be F2.

Table 17. *Acoustic variables predicting the words within the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b). The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

(a)	normal speech		clear speech (SF)		clear speech (HG)	
	subj 1	subj 2	subj 1	subj 2	subj 1	subj 2
V duration	73	73	83	85	77	75
F1	73	68	80	85	89	**93**
F2	**84**	**82**	**100**	**94**	**91**	75
TL	42	41	48	55	**91**	75
f0 rise size	39	50	46	42	51	54
f0 fall size	62	64	52	52	51	50
V duration	73	73	83	85	77	75

(b)	normal speech		clear speech (SF)	
	subj 1	subj 2	subj 1	subj 2
V duration	67	61	77	83
F1	59	**66**	73	**88**
F2	**85**	**66**	**82**	67
TL	56	47	59	58
f0 rise size	48	45	64	54
f0 fall size	48	37	68	46
V duration	67	61	77	83

In Table 17b shows that regarding the /ʊ u uː/ triplet F2 is the best predictor for subject 1 both in normal speech and clear speech. For subject 2 F1 or F2 are the best predictors for normal speech, whereas in clear speech only F1 serves as the best predictor.

Considering all four speakers using all available samples we find that for the /ʊ u uː/ triplet either F1 or F2 has the highest percentage of correctly predicted triplet vowels, as shown in Table 18.

Table 18. *Acoustic variables predicting the words within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly guessed. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

	normal speech				clear speech (SF)				clear speech (HG)			
	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
V duration	60	67	44	60	85	71	68	75	92	74	56	76
F1	63	67	**71**	**67**	81	**92**	**71**	75	67	72	74	**85**
F2	77	**71**	60	63	**92**	77	54	**89**	**88**	**89**	**81**	76
TL	38	40	40	33	46	54	51	39	65	38	39	54

For both triplets we find that spectral features best predict the vowels within a triplet. For the /ɪ i iː/ triplet this is mainly F2, for the /ʊ u uː/ triplet this is either F1 or F2. For f0 we found relative small percentages.

3.3.3 Prediction of triplet words by combinations of variables

In the previous section we found that the lax/tense distinction is best predicted by spectral features, and the short/long distinction by vowel duration. We now look for variables and combinations of two variables that predict all three words within a triplet in normal speech and in clear speech. Again, linear discriminant analysis was used.

Table 19 shows combinations of two acoustic variables predicting the words within the /ɪ i iː/ triplet. We find that V duration + F1 and/or V duration + F2 has the highest percentages across all speakers. This result agrees with the finding in the previous section, where we found that spectral features and vowel duration are the most important predictors.

For subject 1 F2 is sufficient to predict the triplet words with 100% accuracy for clear speech using Saterland Frisian stimuli (see Table 17a). Adding any other variable does not improve or worsen this percentage.

Table 19. *Combinations of two acoustic variables predicting the words within the /ɪ i iː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

		normal speech		clear speech (SF)		clear speech (HG)	
		subject 1	subject 2	subject 1	subject 2	subject 1	subject 2
V duration	F1	88	**91**	98	**100**	98	**100**
V duration	F2	**92**	**91**	**100**	97	**100**	**100**
V duration	TL	69	75	85	85	83	75
V duration	f0 rise size	72	68	87	88	83	89
V duration	f0 fall size	76	80	89	85	85	75
F1	F2	82	82	**100**	**100**	91	89
F1	TL	72	70	85	79	94	93
F1	f0 rise size	70	73	87	88	87	89
F1	f0 fall size	76	80	80	85	91	86
F2	TL	86	84	**100**	97	91	68
F2	f0 rise size	85	84	**100**	97	91	71
F2	f0 fall size	86	80	**100**	97	96	82
TL	f0 rise size	46	45	52	58	62	54
TL	f0 fall size	64	64	61	63	57	57
f0 rise size	f0 fall size	61	68	57	58	68	71

Results for the /ʊ u uː/ triplet are given in Table 20. No consensus is found for the best combination of predicting variables, neither for normal speech nor for clear speech. For normal speech for subject 1 F2+F1 gives the best results, for subject 2 F1+TL yields the highest percentage. Looking at the results for clear speech, for subject 1 F2+f0 fall size and for subject 2 F1or F2 + V duration gives the highest percentage. The lack of consensus for normal speech may be explained by the fact that there is no consensus about the distinction of short and long tense vowels (see Section 2.6.1). For clear speech the result for subject 2 is in line with the previous results. The result for subject 1 is more unexpected. Looking at Table 14, we find that the lax/tense distinction is best predicted by F1 or f0 fall size, and the short/long distinction by V duration, F2 or f0 dynamics. In view of these results, in the combination F2+f0 fall size, f0 fall size is likely the cue for distinguishing short lax and tense vowels, and F2 for distinguishing short and long tense vowels.

Table 20. *Combinations of two acoustic variables predicting the words within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. The highest percentage(s) per subject and per speech type are represented in bold.*

		normal speech		clear speech (SF)	
		subject 1	subject 2	subject 1	subject 2
V duration	F1	78	68	86	**95**
V duration	F2	85	66	86	**95**
V duration	TL	74	71	86	88
V duration	f0 rise size	70	68	82	83
V duration	f0 fall size	74	68	86	92
F1	F2	**89**	68	82	67
F1	TL	78	**76**	77	79
F1	f0 rise size	78	66	77	79
F1	f0 fall size	74	68	77	88
F2	TL	85	74	82	71
F2	f0 rise size	85	68	82	67
F2	f0 fall size	85	68	**91**	67
TL	f0 rise size	63	58	64	67
TL	f0 fall size	67	42	68	67
f0 rise size	f0 fall size	56	37	68	50

Results obtained on the basis of four speakers are shown in Table 21. We do not find any consensus across speakers for normal speech, since no consensus could be found for the short/long distinction (see Section 2.6.1). Considering the results for clear speech, however, the combination of V duration + F1 and/or F2 gives the highest percentage of correctly predicted triplet words, except for clear speech pronounced by subject 3 when using High German stimuli, in which case the combination F1+F2 predicts triplet words the best.

Table 21. *Combinations of two acoustic variables predicting the words within the /ʊ u uː/ triplet. The numbers represent the percentage of triplet words that were correctly estimated. Variables regarding f0 are not considered. The highest percentage(s) per subject and per speech type are represented in bold.*

		normal speech				clear speech (SF)				clear speech (HG)			
		subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4	subj 1	subj 2	subj 3	subj 4
V duration	F1	77	67	75	63	**92**	**98**	**88**	**97**	94	**100**	81	**96**
V duration	F2	72	71	65	63	92	94	**88**	97	**96**	96	78	89
V duration	TL	59	73	46	**69**	79	85	78	75	92	74	69	76
F1	F2	**78**	65	77	69	**92**	81	70	83	94	96	**89**	89
F1	TL	70	73	69	60	81	88	76	83	69	77	70	81
F2	TL	72	**75**	58	60	**92**	81	59	86	90	89	81	81

Note that the results for the /ʊ u uː/ triplet on the basis of the four speakers show a much clearer pattern than the results on the basis of the two speakers. The results on the basis of the four speakers may be considered more reliable since they are obtained on the basis of a larger number of samples per speaker (see Table 9) apart from the fact that the number of speakers is larger. Taking this into account we conclude again that the distinction of the triplet words of the /ɪ i iː/ triplet and the /ʊ u uː/ is best predicted by a combination of a spectral features (F1 or F2) and vowel duration.

When comparing the results of this section with the results for individual variables in Section 3.3.2, we find that the percentages on the basis of two variables are slightly higher than the percentages obtained on the basis of one variable. The differences between clear speech and normal speech are much bigger, whereas differences between clear speech (SF) and clear speech (HG) are negligible.

3.4 Normal speech versus clear speech

In this section we study in more detail the effect clear speech has on duration, spectral properties, and f0 compared to normal speech. Figure 9 shows durations of speech segments within the /ɪ i iː/ triplet words for subject 1 and subject 2, shown as percentages of the total word length. For both speakers we find that the long tense vowel in Smíete has become proportionally longer. For all triplet words and for both speakers we find that the final /ə/ has proportionally been shortened in clear speech. All of these differences are significant at the 0.05 level.

Considering the total word duration, the triplet words Smitte and smiete are longer in normal speech than in clear speech. The word Smíete is longer only for subject 2.

We do not find any consensus for the two subjects regarding spectral differences, neither do we find consensus regarding f0 differences, except that f0 fall size in the short lax vowel in Smitte is larger in clear speech than in normal speech.

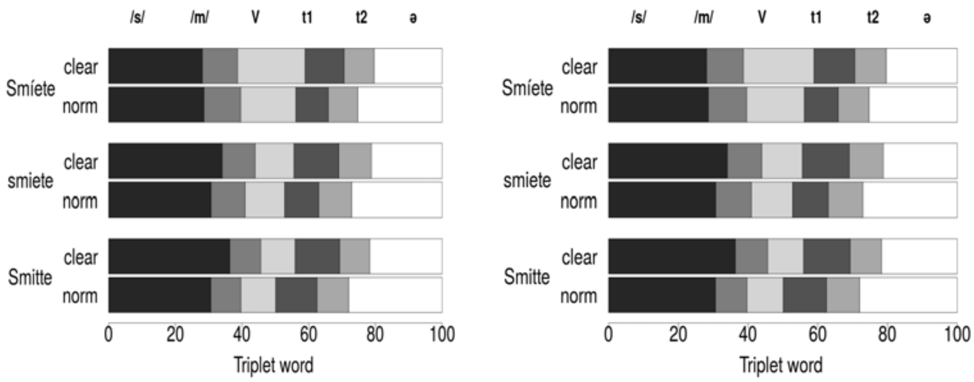


Figure 14. *Proportional durations of speech segments within the /ɪ i iː/ triplet words for subject 1 (left) and subject 2 (right)l shown as percentages of the total word duration.*

Figure 10 shows durations of speech segments within the /ʊ u uː/ triplet words for subject 1 and subject 2, shown as percentages of the total word length. We find that clear speech does not affect the proportional durations of speech segments and the total word durations of subject 1. For subject 2 clear speech causes an increased proportional duration of initial /f/ and a decreased proportional vowel duration for all triplet words. The final /l/ in ful is proportionally shortened as well as the total duration of the triplet word fuul.

There are no consensus differences in proportional duration and in spectral properties when comparing clear speech with normal speech. However, when regarding f0, for both subjects we found that f0 fall size and f0 dynamics in the short tense vowel in fuul increases in clear speech.

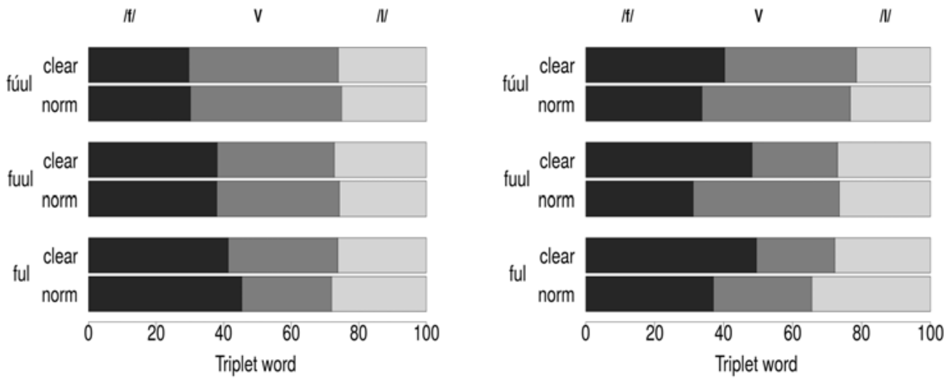


Figure 15. *Proportional durations of speech segments within the /ʊ u uː/ triplet words for subject 1 (left) and subject 2 (right), shown as percentages of the total word length.*

When considering durations of speech segments within the /ʊ u uː/ triplet words for all four subjects on the basis of samples which do not necessarily have a clear f0 peak, for subject 2 we again find that clear speech causes an increased proportional duration of initial /f/ and a decreased proportional vowel duration for all triplet words. For subject 4 we find similar results: proportional duration of /f/ in *fuul* and *fúul* is increased in clear speech and the proportional vowel duration in *fuul* is decreased. For subject 3 we find that the proportional vowel duration of the long tense vowel in *fúul* is increased in clear speech. Subject 4 pronounces the final /l/ in *fúul* relatively longer, subject 2 does the same both in *ful* and *fúul*.

There are no consensus differences in proportional duration and in spectral properties when comparing clear speech with normal speech. We find a tendency of the total word duration to be shortened in clear speech for all triplet words. For the triplet word *ful* this is the case for subjects 1, 2, and 3, for the triplet word *fuul* this is the case for subjects 1, 2, and 4, for the triplet word *fúul* this is the case for subject 4 only. Regarding the spectral properties, for subject 1, subject 2, and subject 4 we found that the short tense vowel /u/ in *fuul* has higher F1 and F2 frequencies in clear speech.

When comparing the effect of clear speech on the triplets /ɪ i iː/ and /ʊ u uː/, we do not find any consistent effects for both triplets. High German translations of the words of the /ʊ u uː/ triplet differ from the Saterland Frisian words only in their nuclei, which may have led the subjects to pronounce the words even more clearly. The German translations of two words of the /ɪ i iː/ triplet are not cognates of the Saterland Frisian words, which might have made the subjects less confident.

Within the /ɪ i iː/ triplet there is consensus among the speakers for proportionally lengthening the long tense vowel in *Smíete* and shortening the final schwa in clear speech. Within the /ʊ u uː/ triplet we do not find any consensus effect regarding proportional duration and spectral properties for the speakers. However, we found that both subject 1 and 2 increase f0 fall size and f0 dynamics in the short tense vowel in fuul in clear speech.

Regarding f0 differences, for both triplets and for both subject 1 and subject 2 we find a higher f0 peak for short tense vowels in clear speech. Subject 1 also has a higher f0 peak for the long tense vowel in the /ɪ i iː/ triplet.

4. Experiment 3: Perception

4.1 Procedure

In this experiment we tested to what extent the words were perceived as intended by the speaker. All four subjects listened to the triplet words. As stimuli we used the realizations recorded from subject 1 in the clear speech experiment (SF), as subject 1 was the most confident speaker and only in clear speech all vowels were distinguished within both triplets (see Table 11).

For each word the participants indicated which word within the triplet was pronounced. There were two tests, one testing the *Smitte/smiete/Smíete* triplet and one testing the *ful/fuul/fúul* triplet. The first test was carried out by subject 1 and subject 2 only, since the other subjects were not familiar with all of the words of the *Smitte/smiete/Smíete* triplet. The second test was carried out by all four subjects. The stimuli were presented via headphones.

Each test started by a short practice session, in which six triplet words were presented in a randomized order. After this session, three blocks followed. Since each of three triplet words was represented by 15 instances, in each block 45 triplet words were presented.

When a word was played, the triplet words were presented on the screen. Since the orthography of Saterland Frisian would suggest an order in duration due to the use of diacritics and/or double vowels, High German translations were shown on the screen instead (see Figure 16, left picture). A choice was made by clicking on the line close to the word the subject thought she heard. A little triangle showed the choice and appeared when the mouse pointer was moved on top of the line (see Figure 16, right picture).

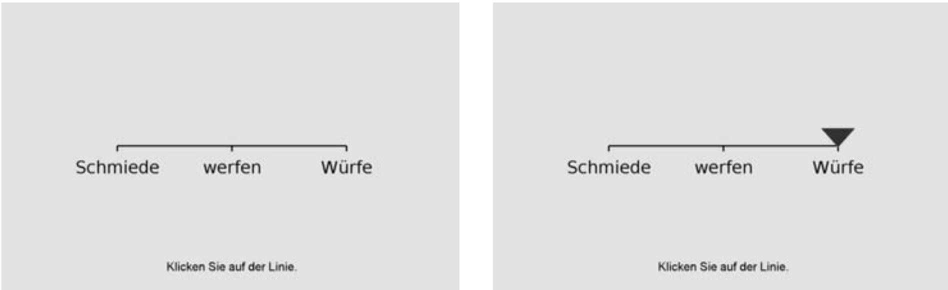


Figure 16. *Screenshots of the perception experiment.*

4.2 Agreement between production and perception

Results for the /ɪ i iː/ triplet are shown in Table 22. The tables show a strong agreement between the stimulus which was intended to be pronounced and the triplet word perceived by the speakers. For subject 2 there is a perfect agreement, for subject 1 only 2% of the words intended to be pronounced with the short tense /i/ (smiete) were misperceived as words with the long tense /iː/.

Table 22. *For the intended production of each word within the /ɪ i iː/ triplet the percentages of word forms perceived by subject 1 (left) and subject 2 (right) are given.*

		perceived by subject 1		
		Smitte	smiete	Smíete
stimulus	Smitte	100	0	0
	smiete	0	98	2
	Smíete	0	0	100

		perceived by subject 2		
		Smitte	smiete	Smíete
stimulus	Smitte	**100**	0	0
	smiete	0	**100**	0
	Smíete	0	0	**100**

Table 23 shows the agreement between production and perception for the /ʊ u uː/ triplet for all four speakers. Again, there is a strong agreement, albeit slightly less strongly than for the /ɪ i iː/ triplet. Regarding the confusions, *ful* is mostly misperceived as *fúul*, which is unexpected. We would rather have expected that ful would be confused with *fuul*, both forms having a short vowel. This unexpected pattern is consistent across all speakers and hence suggests that the confusion does not lie in perception but might be rather due to the fact, that the stimulus was mistaken as *fúul* and pronounced respectively when originally recorded. On average, *fuul* is most frequently confused with *fúul*, and *fúul* with *fuul*, both words having a tense vowel.

***Table 23**. For the intended production of each word within the /ʊ u uː/ triplet the percentages of word forms perceived by subject 1 (top left), subject 2 (top right), subject 3 (bottom left) and subject 4 (bottom right) are given.*

		perceived by subject 1		
		ful	fuul	fúul
stimulus	ful	**93**	0	7
	fuul	0	**98**	2
	fúul	0	9	**91**

		perceived by subject 2		
		ful	fuul	fúul
stimulus	ful	**93**	0	7
	fuul	0	**93**	7
	fúul	0	7	**93**

		perceived by subject 3		
		ful	fuul	fúul
stimulus	ful	**93**	0	7
	fuul	0	**84**	16
	fúul	0	4	**96**

		perceived by subject 4		
		ful	fuul	fúul
stimulus	ful	**91**	2	7
	fuul	7	**89**	4
	fúul	0	18	**82**

We conclude that for both triplets there is a strong agreement between the intended pronunciation of the triplet words and the way in which they were perceived by the speakers.

4.3 Cues predicting triplets

In Sections 2.6 and 3.3 we found that duration and spectral features predict the words within a triplet best. In these analyses we used linear discriminant analysis in order to predict the triplet words which were intended to be pronounced. In this section we will consider duration, F1 and F2 and find how well they predict the word forms as they were perceived by the speakers. Results will be compared to the percentages of correctly predicted intended word forms.

In Table 24a the percentages of correctly predicted intended triplet words and the triplet words perceived by the two speakers are shown for the /ɪ i iː/ triplet. Percentages are high and all of them significant with $p<0.0001$. F2 only yields the highest percentages, adding another second variable will not improve the prediction. Results for production and perception are comparable: somewhat lower percentages for vowel duration and F1 only (84%-85%), and higher percentages for the other variables or combinations of variables (97%-100%).

Table 24. *Prediction of the intended production and the perceived word (subject 1 and subject 2) by (combinations of) duration, F1, and F2 for the /ɪ i iː/ triplet (a) and the /ʊ u uː/ triplet (b).*

(a)			prod.	perc.	perc.
			subj 1	subj 1	subj 2
dur			84%	84%	84%
F1			84%	85%	84%
F2			100%	99%	100%
dur	F1		98%	97%	98%
dur	F2		100%	99%	100%
F1	F2		100%	99%	100%
dur	F1	F2	100%	99%	100%
dur			84%	84%	84%

(b)			prod.	perc.	perc.	perc.	perc.
			subj 1	subj 1	subj 2	subj 3	subj 4
dur			84%	87%	87%	86%	79%
F1			82%	84%	83%	81%	82%
F2			91%	94%	95%	90%	86%
dur	F1		91%	96%	96%	94%	90%
dur	F2		91%	96%	96%	94%	92%
F1	F2		91%	94%	95%	90%	84%
dur	F1	F2	91%	96%	96%	94%	90%
dur			84%	87%	87%	86%	79%

Results for the /ʊ u uː/ triplet are shown in Table 24b. All percentages are high and significant with $p<0.0001$. The combination of duration and F2 gives the highest percentages for all subjects. Again, we find a high agreement between production and perception, with lower percentages for vowel duration and F1 individually (79%-87%) and higher percentages for other variables or combinations of variables (84%-96%).

When considering individual predictors, we find for both triplets that F2 predicts both the intended pronunciation of the triplet words and the way they were perceived by the speakers best. For the /ɪ i iː/ triplet adding other variables – F1 and duration – does not improve the percentage of correctly predicted triplet words. For the /ʊ u uː/ triplet, however, adding duration improves the percentages. Both for production and perception high percentages are obtained, varying between 86% and 100% when using F2 only.

5. Conclusions and discussion

In this paper we presented two production experiments and one perception experiment, by means of which we systematically searched for the cues, which distinguish the vowels being the nuclei in the triplet Smitte / smiete / Smíete and ful / fuul / fúul. The two production experiments are a normal speech experiment and a clear speech experiment.

Normal speech

In the normal speech experiment we found for both triplets that short lax and tense vowels were distinguished by F1 and F2. In the /ɪ i iː/ triplet short and long tense vowels were distinguished by vowel duration and F2, but in the /ʊ u uː/ triplet this distinction was not found for any variable. The distinction between short lax and short tense vowels was predicted by spectral variables, F1 and/or F2 for all subjects and both triplets. Within the /ɪ i iː/ triplet short and long tense vowels were predicted by vowel duration. For the /ʊ u uː/ triplet we could not find a variable that predicts the distinction between short and long tense vowels best for all speakers. All three vowels in the /ɪ i iː/ triplet were predicted best by a combination of F2 and vowel duration, but for the /ʊ u uː/ triplet a consensus combination could not be found.

As to the role of vowel duration, we found that all three triplet words of the /ɪ i iː/ triplet are distinguished from each other by vowel duration. For the /ʊ u uː/ triplet short lax and tense vowels are distinguished by duration when considering the consensus of subjects 1 and 2, but this consensus cannot be found when considering all four subjects.

Clear speech

The main finding from the clear speech condition is that short and long tense vowels in the /ʊ u uː/ triplet were distinguished. The distinguishing cues may differ depending on the data set we used but we found consensus about vowel duration distinguishing short and long tense vowels. Eventually, for both triplets, short and long tense vowels are distinguished by vowel duration. Additionally, we found that vowel duration has the highest percentage of correctly predicted short tense and long tense vowels. The distinction of the triplet words of the /ɪ i iː/ triplet and the /ʊ u uː/ triplet is best predicted by a combination of spectral features (F1 or F2) and vowel duration. Our results are similar to Kohler (2001) for 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 short lax and tense vowels within a triplet differ qualitatively, and short and long tense vowels differ quantitatively.

Considering the role of vowel duration for the /ɪ i iː/ triplet, all three triplet words are distinguished from each other by vowel duration. For the /ʊ u uː/ triplet only short and long tense vowels are distinguished by duration.

Siebs' (1889) mentioned the use of Stoßton and Schleifton in Saterland Frisian. This suggests that f0 might play an additional role. Among others, Yu (2010) found that listeners perceived stimuli with f0 dynamics to be longer even when the stimuli have the same acoustic duration. We also studied the role of f0, and for the /ɪ i iː/ triplet we found that in normal speech as well as clear speech (HG) long tense vowels have a larger f0 fall size than short lax vowels. In clear speech we find for both triplets that short tense and long tense vowels are in addition to vowel duration (and to other variables in case of the /ɪ i iː/) distinguished by f0 dynamics. Short tense vowels have significantly more f0 dynamics than long tense vowels. These results indicate that f0 plays a role in the distinction of short and long tense vowels.

The latter result is unexpected. For example, Gordon (2001) mentions the hierarchy of tone bearing ability, in which syllables with a greater overall sonorous energy are better equipped to support contour tones than syllables with lesser sonorous energy. Yu (2010) mentions that "it is repeatedly observed that dynamic tones tend to be restricted to phonetically long sonorous segments" and refers to Gordon (2001) and others. In the light of this we would have expected that the long tense vowels would have more f0 dynamics than the short tense vowels, but our results are the other way around. The short tense vowels with most f0 dynamics make them perceptually longer and therefore closer to the long tense vowels, and more distant to the short lax vowels. With reference to Figures 3 and 6 it may be noted that the durations of short tense vowels are on average closer to the durations of short lax vowels than to the durations of the long tense vowels. For the /ʊ u uː/ triplet short lax and tense vowels are not significantly distinguished by duration in clear speech (see Tables 11 and 12). It is thus likely that f0 dynamics serves as an enhancing factor to make short tense vowels perceptually more distinct from short lax vowels.

Several studies such as Smiljanić & Bradlow (2005) compared clear speech to normal speech. According to those studies, clear speech may decrease the speaking rate, which involves longer segments, an expansion of the vowel space, and an increase in the pitch range. We compared the effect of clear speech on the triplets /ɪ i iː/ and /ʊ u uː/. For both triplets and for both subject 1 and subject 2 we found a higher f0 peak for short tense vowels in clear speech. Subject 1 also had a higher f0 peak for the long tense vowel in the /ɪ i iː/ triplet. For the other variables we did not find consistent effects for both triplets and all subjects.

Within the /ɪ i iː/ triplet we found consensus among the speakers for proportional lengthening of the long tense vowel in Smíete and shortening of the final schwa in clear speech. Within the /ʊ u uː/ triplet we did not find any consensus effect regarding proportional duration and spectral properties for all speakers. However, we found that both subject 1 and 2 increase f0 fall size and f0 dynamics in the short tense vowel in fuul in clear speech.

Perception

The production experiments suggest that the lax/tense distinction is explained by spectral differences (F1 and/or F2), while the short/long distinction is explained by duration (and F2). In addition to the production experiments where we tried to explain the intended pronunciation of samples on the basis of acoustic variables that were measured on the basis of those samples, we carried out a perception experiment where subjects listened to word samples which enabled us to explain the perceived pronunciation by acoustic measurements that were obtained on the basis of the samples the subjects heard.

In section 1 we asked whether native speakers of Saterland Frisian still perceive the phonological distinctions within a triplet. For both triplets we find a strong agreement between the intended pronunciation of the triplet words and the way in which they were perceived by the speakers.

When considering individual predictors, we find for both triplets that F2 predicts both the intended pronunciation of the triplet words and the way they were perceived by the speakers best. For the /ɪ i iː/ triplet adding other variables – F1 and duration – does not improve the percentage of correctly predicted triplet words. For the /ʊ u uː/ triplet, however, adding duration improves the percentages. The latter case is in agreement with the results of the production experiments that suggest that a combination of spectral features and duration explains the distinctions within a triplet. For the /ɪ i iː/ triplet, however, it is remarkable that combination of duration and F2 does not improve the prediction that is made on the basis of only F2. However, we have to keep in mind that the perception test is based on the clear speech samples of subject 1 only. Considering the vowel plot of subject 1 obtained on the basis of the clear speech samples of the /ɪ i iː/ triplet (SF version) in Figure 4, we find that all three ellipses are clearly distinguished in the F2 dimension.

Discussion

Our finding that short lax and tense vowels are spectrally distinguished and short and long tense vowels primarily are distinguished by duration is consistent with the representation of Saterland Frisian vowel categories by Kramer (1968) and Fort (1980, p. 62). The high vowels /i iː/ and /u uː/ are only distinguished by a length mark, suggesting a durational difference, whereas /ɪ/ and /ʊ/ are classified among the close-mid vowels as short counterparts of /eː/ and /oː/, which suggests a spectral difference compared to /i/ and /u/ respectively.

Our results agree with the study of de Graaf (1985), who studied the distinction between high vowels /i iː/, /y yː/ and /u uː/ for West Frisian. The durational difference between long and short high tense vowels was significant – which agrees with our findings. Similar results were found by Gilbers, Visser & Weening (2012, p. 69-90).

Regarding the spectral properties, de Graaf found that all speakers pronounced the short vowels in a more central position than the long vowels. The more central position of the short vowels agrees partly with our results. We found that /i/ has a larger F1 (in clear speech) and a smaller F2 than /iː/ (in normal speech and in clear speech using Saterland Frisian stimuli). The centralization of /u/ compared to /uː/ is only found in the results of the clear speech experiment carried out by four subjects using High German stimuli.

The work presented in this paper should be considered as an exploratory study. We notice that explanatory factors for the distinctions between vowels were not always (consistently) identified. For normal speech, the triplet words fuul and fúul were not distinguished by any acoustic features. Other or stronger acoustic features than the ones considered in this paper might add to the distinction of these vowels. F3 and F4 were not considered and this may be a problem when finding the cues that identify the difference between the tense and lax vowels of the triplets, whether they are front or back.

It has been shown in various publications (cf. Schwartz et al. 1997) that both dispersion and focalisation factors should be considered when investigating the distinction of vowels. The Dispersion-Focalization Theory (DFT) attempts to predict vowel systems based on the minimization of an energy function summing two perceptual components: global dispersion, which is based on inter-vowel distances; and local focalization, which is based on intra-vowel spectral salience related to the proximity of formants (Schwartz et al. 1997). Particularly important is the focalization aspect. Focalization is the convergence between two consecutive formants in a vowel spectrum.

One important feature which we would like to include is effective F2, which describes in a non-linear way the combining effect of F2 and higher formants (Carlson , Granström & Fant 1970). For example, Fant (1973:52) found that Swedish back vowels can be approximated well using only F1 and F2, whereas for some other vowels, especially front-rounded vowels, F3 and higher formants should also be considered. Fant formulated a value located between F2 and F3 as follows:

$$F2 = F2 + 1/2(F3 – F2) * (F2 – F1) / (F3 - F2)$$

Mapping the vowels in the F1 and effective F2 space more successfully separated the vowels, that overlap heavily in the natural F1/F2 plane. Kmoshita (2000) found a better separation between Japanese back vowels /o/ and /u/, and /o/ and /a/ when using effective F2 instead of F2. It seems, therefore, important to identify effective F2 and not only to rely on F2 values to be opposed to F1. Leaving aside the duration of tense and lax vowels, this could characterize in a better way the pairs /ɪ-i/ and /ʊ-u/ in our study. Schwartz et al. (1997) computed effective F2 from F2, F3, and F4 using a model based on the concepts of the center of gravity and the 3.5-Bark critical distance dc introduced by Chistovich & Lublinskaya (1979).

Therefore, including effective F2 in the analysis and considering both dispersion and focalisation factors would be useful future work that will extend and deepen our understanding of Saterland Frisian vowel variation even more.

Funding

The research reported in this paper has been funded by the Deutsche Forschungsgemeinschaft (DFG), grant number PE 793/2-1.

Acknowledgements

We would like to thank the informants for participating in our study. We are grateful to Darja Appelganz, Romina Bergmann, Dorothee Lenartz, Michaela Ballin, and Nicole Lommel for labeling the recordings in PRAAT. We thank an anonymous reviewer for his valuable comments.

**Fryske Akademy, The Netherlands*
***Institute of German Studies, University of Oldenburg, Germany*

REFERENCES

Boersma, P. & Weenink, D. (1992-2015). Praat: Doing Phonetics by Computer. Available at http://www.praat.org.
Bradlow, A.R. (2002). Confluent talker- and listener-related forces in clear speech production. In C. Gussenhoven & N. Warner (Eds.), Laboratory Phonology 7 (pp. 241-273). Berlin: Mouton de Gruyter.
Bradlow, A.R., Kraus, N. & Erin, H. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. Journal of Speech, Language, and Hearing Research, 46(1), 80-97.
Carlson, R., Granström, B. & Fant, G. (1970) Some studies concerning perception of isolated vowels. STL - QPSR , 2-3, 19-35.
Chistovich, L.A. & Lublinskaya, V.V. (1979). The center of gravity ef fect in vowel spectra and critical distance between the formants. Hearing Research, 1, 185-195.
Cumming, R.E. (2011), The effects of dynamic fundamental frequency on the perception of duration. Journal of Phonetics, 39, 375-387.
Fant, G. (1973). Speech Sound and Features. Cambridge: MIT Press.
Ferguson, S.H. & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America 112, 259-271.
Fort, M.C. (1980). Saterfriesisches Wörterbuch mit einer grammatischen Ubersicht. Hamburg: Buske Verlag. In collaboration with Hermann Dumstorf.
Fort, M.C. (2015). Saterfriesisches Wörterbuch. Mit einer phonologischen und grammatischen Übersicht. Hamburg: Buske Verlag.
Fox, R.A. & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. The Journal of the Acoustical Society of America, 126, 2603-2618.
Gagne, J.-P., Rochette, A.-J. & Charest, M. (2002). Auditory, visual, and audiovisual clear speech. Speech Communication, 37, 213-230.
Gilbers, D.G., Visser, W. & Weening, J. (2012). Het verschil tussen de lange en korte hoge vocalen in het Fries en de invloed hiervan op de realisering van de Nederlandse hoge vocalen door Friestaligen. Us Wurk, 61, 69-90.
Gordon, M. (2001), A typology of contour tone restriction. Studies in Language, 25(3), 423-462.
Graaf, T. de (1985). Phonetic aspects of the Frisian vowel system. NOWELE, 5, 23-40.

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

Gussenhoven, C. (2007). Vowel height split explained: Compensatory listening and speaker control. In J. Cole & J.I. Hualde (Ed.), Laboratory Phonology 9 (pp. 145-172), Berlin: Mouton de Gruyter.

Hastings, C., Mosteller, F., Tukey, J.W. & Winsor, C.P. (1947). Low moments for small samples: a comparative study of order statistics. Annals of Mathematical Statistics, 18, 413-426.

Heeringa, W., Peters, J. & Schoormann, H. (2014). Segmental and prosodic cues to vowel identification: The case of /ɪ i iː/ and /ʊ u uː/ in Saterland Frisian. In: N. Campbell, D. Gibbon & D. Hirst (Eds.), Social and Linguistic Speech Prosody Proceedings of the 7th international conference on Speech Prosody (pp. 643-647). Dublin: Science Foundation Ireland.

Jin, S.-H., Liu, C. (2013). The vowel inherent spectral change of English vowels spoken by native and non-native speakers. The Journal of the Acoustical Society of America, 133(5), 363-369.

Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. Language, 69, 505-528.

Kinoshita, Y. (2000). Effective F2 as a parameter of Japanese forensic speaker identification. Proceedings of the 8th International Conference on Speech Science and Technology, (pp. 160-165). Canberra, Australia.

Kohler, K.J. (2001). Uberlänge im Niederdeutschen? In R. Peters, H.P. Pütz & U. Weber (Eds.), Vulpis Adolatio. Festschrift für Hubertus Menke zum 60. Geburtstag (pp. 385-402). Heidelberg: C. Winter.

Kramer, P. (1968), De Sealter lûden, Us Wurk, 17, 21-24.

Kramer, P. (1982). Kute Seelter Sproakleere – Kurze Grammatik des Saterfriesischen. Rhauderfehn: Ostendorp.

Krause, J.C. & Braida, L.D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. The Journal of the Acoustical Society of America, 115, 362-378.

Lehiste, I. (1976). Influence of fundamental frequency pattern on the perception of duration, Journal of Phonetics, 4, 113-117.

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

Liu, S., Del Rio, E., Bradlow, A.R. & Zeng, F.G. (2004). Clear speech perception in acoustic and electrical hearing. The Journal of the Acoustical Society of America, 116, 2374-2383.

Mayr, R. & Davies, H. (2011). A cross-dialectal acoustic study of the monophthongs and diphthongs of Welsh. Journal of the International Phonetic Association, 41, 1-25.

Moon, S.-J. & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. The Journal of the Acoustical Society of America, 96, 40-55.

Peters, J. (2008). Saterfrisian intonation. An analysis of historical recordings. Us Wurk, 57, 141-169.

Picheny, M.A., Durlach, N.I. & Braida, L.D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. Journal of Speech and Hearing Research, 29, 434-446.

Picheny, M.A., Durlach, N.I. & Braida, L.D. (1989). Speaking clearly for the hard of hearing. III. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. Journal of Speech and Hearing Research, 32, 600-603.

Prehn, M. (2007). Schwa loss and its results in LG. Tone or Overlength? In B. Los & M. van Koppen (Eds.), Linguistics in the Netherlands 2007, AVT Publications 24, 187-198. Amsterdam: John Benjamins.

Prehn, M. (2012). Vowel quantity and the fortis – lenis distinction in North Low Saxon. PhD Thesis University of Amsterdam, LOT 291.

Rogers, C.L., DeMasi, T.M., Krause, J.C. (2010). Conversational and clear speech intelligibility of ⁄bVd⁄ syllables produced by native and non-native English speakers. The Journal of the Acoustical Society of America, 128(1): 410-423.

Schwartz, J.L., Boë, L.J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of vowel systems. Journal of phonetics, 25(3), 255-286.

Siebs, Th. (1889). Zur Geschichte der englisch-friesischen Sprache. Halle (Saale): Max Niemeyer.

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

Smiljanić, R. & Bradlow, A.R (2005). Production and perception of clear speech in Croatian and English. Journal of the Acoustical Society of America, 118(3), 1677-1688.

Stellmacher, D. (1998). Das Saterland und das Saterländische. Oldenburg: Florian Isensee.

Tiersma, P.M. (1999). Frisian Reference Grammar. Ljouwert: Fryske Akademy. 2nd edition.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. The Journal of the Acoustical Society of America, 88, 97-91.

Traunmüller, H. & Eriksson, A. (1994). The size of F0-excursions in speech production and perception. Working Papers, 43, 136-139. Department of Linguistics and Phonetics, Lund University.
Tröster-Mutz, S. (1997). Phonologie des Saterfriesischen. MA Thesis University of Osnabrück.
Tröster-Mutz, S. (2002). Untersuchungen zu Silbenschnitt und Vokallänge im Saterfriesischen. Theorie des Lexikons, Arbeiten des SFB 282, 120, 1-27. Düsseldorf: Heinrich-Heine-University.
Uchanski, R.M. (1988). Spectral and temporal contributions to speech clarity for hearing impaired listeners. Unpublished doctoral dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.
Uchanski, R.M. (1992). Segment durations in conversational and clear speech. Unpublished manuscript.
Venables, W.N. & Ripley, B.D. (2002). Modern Applied Statistics with S. Fourth edition. New York: Springer.
Visser, W. (2003). Lange en koarte ienlûden yn it Frysk: in reaksje op De Haan (1999). Us Wurk, 52, 130-154.
Wilcox, R.R. (2012). Modern Statistics for the Social and Behavioral Sciences: A Practical Introduction. London: Chapman & Hall/CRC.
Yu, A.C.L. (2010). Tonal effects on perceived vowel duration. In C. Fougeron, B. Kuehnert, M. D'Imperio & N. Vallé (Eds.), Laboratory Phonology 10 (pp. 151-168). Berlin: Mouton de Gruyter.