Voicing alterations in Icelandic sonorants – a photoglottographic and acoustic analysis

Lasse Bombien

Institut für Phonetik und digitale Sprachverabeitung, Christian-Albrechts-Universität Kiel, Germany

---

This study aims to provide experimental data for the phonatory behaviour during the production of the Icelandic voiceless sonorants [l̥ n̥ r̥] in contrast with their voiced counterparts in word initial and medial position. Based on the analysis of photoglottographic and acoustic data of two Icelandic speakers, sonorants in Icelandic are shown not to differ by means of a classic "voiced vs. voiceless" contrast. Instead, the phonation type of the "voiceless" sonorants varies in dependence of context and manner of articulation. Furthermore, additional mechanisms of sound production are involved in the distinction of these sounds, such as frication.

---

1 Introduction

Icelandic has a set of so called "voiceless" sonorants which can appear in word medial position preceding stops as well as in initial position followed by vowels. In both positions they appear to contrast with their voiced counterparts, see the examples in (1a) for the initial context and in (1b) for the medial context[1]:

		nýta	[niːta]	(to use)	*hnýta*	[n̥iːta]	(to knot)
(1)	a.	*líða*	[liːða]	(to feel)	*hlíða*	[l̥iːða]	(slope)
		rífa	[riːva]	(to rip)	*hrífa*	[r̥iːva]	(to enchant)
		hendi	[hɛntɪ]	(hand)	*henti*	[hɛn̥tɪ]	‚to dispose of'
	b.	*eldi*	[ɛltɪ]	(to cook)	*elti*	[ɛl̥tɪ]	(to chase)
		björg	[pjørk]	(rescue)	*björk*	[pjør̥k]	(birch)

While considerable effort has gone into phonological systematisation of these sounds, comparatively little work has been done to describe them from an articulatory point of view. Section 2.1 provides an overview of previous works.

This study is inspired by the observation that there may be continuous vocal-cord vibration throughout voiceless Icelandic nasals[2]. However, this glottal activity is very low in amplitude when compared to voiced nasals or vowels and periodicity is not clearly trackable, indicating a rather lax mode of phonation compared to modal voice. This observation suggests that the distinction of the nasals – and maybe other sonorants, as well – might not be a voiced-voiceless distinction in a purely phonetic sense.

Two major questions are investigated in this work:

---

[1] Some Icelandic words used in this study are inflected.

[2] This observation was a byproduct of a preliminary study which aimed to investigate voiced and voiceless nasals by means of electroglottography and pneumotachography.

1 *Are there multiple mechanisms that contribute to the distinction of voiced vs. voiceless sonorants in Icelandic?*

2 *Do the mechanisms vary between the different classes of sonorants?*

To clarify the exact types of phonation involved in the production of voiceless sonorants photoglottographic and acoustic investigations were carried out.

Concerning the additional mechanisms, this study concentrates on duration and frication. Duration has been found to be greater for voiceless sonorants than for voiced (Pétursson 1977; Johannson 2004; Hoole 1987). Investigating frication as a possible mechanism is inspired by the fact that e.g. the Czech language has two contrastive alveolar trills one of which has a raised tongue position allowing for frication (Dankovicová 1999). It would be possible that frication is added in voiceless Icelandic sonorants to enhance their audibilty. To detect frication, the zero crossing rate is used here, as frication causes high frequencies and therefore high zero crossing rates (Weigelt, Sadoff, and Miller 1990).

Allthough there will be evidence that voiceless sonorants do not, in fact, have to be voiceless, the terms voiceless and voiced will be used to mark the contrast investigated here. Accordingly, the subring diacritic will be used for transcriptions, e.g. [n̥]. In figures where no IPA signs could be used, upper case letters will be used for voiceless sonorants in opposition to lower case letters for voiced sonorants.

2 Background

2.1 Phonetics and Phonology

In Icelandic, voiceless sonorants in medial position always precede stops and appear to occur in the same environments as preaspiration does. Therefore, a short overview of the distributions of Icelandic stops will be given here. Furthermore, some remarks will be made regarding the the dialectal differences between *linmæli* spoken in northern Iceland and *harðmæli* spoken in southern Iceland including the capital, Reykjavík.

There are four places of articulation for stop consonants in Icelandic. (2) shows stops in initial position. Both aspirated and unaspirated stops can occur in initial position, Icelandic does not have voiced stops. Thus, the stop phonemes will be refered to as *lenis* vs. *fortis* rather than as *voiced* vs. *voiceless* here. While the bilabial and the dental stop can be followed by any of the Icelandic vowels, palatal stops are confined to front vowel environments and velar stops to mid and back vowel environments. Therefore palatal and velar stops can be regarded as allophones of /ɡ/ and /k/.

(2)	*banna*	[panːa]	(to prohibit)	*panna*	[pʰanːa]	(pan)
	dýna	[tiːna]	(mattress)	*týna*	[tʰiːna]	(to lose)
	gervi	[cɛrvɪ]	(mask)	*kervi*	[cʰɛrvɪ]	(system)
	galdur	[kaltʏr]	(magic)	*kaldur*	[kʰaltʏr]	(cold)

As shown in example (3), there is no difference between the two dialects concerning initial plosives. However, for stops in medial or final position, the dialects differ. (3) shows how fortis stops differ in aspiration in medial (a) and final (b) position and in clusters (c) after long vowels (Botma 2004, 228). Due to a weakening process, lenis stops are realized as voiced fricatives or approximants in intervocalic or final position.

			northern Icelandic	southern Icelandic	
(3)	a.	*dýpi*	[tiːpʰɪ]	[tiːpɪ]	(depth)
	b.	*sök*	[søːkʰ]	[søːk]	(charge)
	c.	*nepja*	[nɛːpʰja]	[nɛːpja]	(chilliness)
		vökva	[vøːkʰva]	[vøːkva]	(to water)
		depra	[tɛːpʰra]	[tɛːpra]	(dimness)

In medial and final position after short vowels, long aspirated and unaspirated stops can occur, see (4). Aspiration is here realized as *preaspiration*.

(4)	*breiddi*	[prɛɪtːɪ]	(to spread, 1.sg. past)
	skegg	[skɛkː]	(beard)
	breytti	[prɛɪhtɪ]	(to change, 1.sg. past)
	breytt	[prɛɪht]	(to change, past part.)

Preaspiration can also occur alternatingly (5a) and before clusters of stop + /l/ or /n/ (5b). The examples are taken from Thráinsson (1978, 7 f.) and Botma (2004, 230). With regard to preaspiration *linmæli* and *harðmæli* do not differ

		fem. sg.			neut. sg.	
(5)	a.	*feit*	[fɛɪːt]	(fat)	*feitt*	[fɛɪht]
		sæt	[saɪːt]	(sweet)	*sætt*	[saɪht]
	b.	*epli*	[ɛhplɪ]	(apple)		
		læknir	[laɪhknɪr̥]	(physician)		

Voiceless sonorants occur word-medially after short vowels before fortis stops which is the same environment as for preaspiration. (6) shows voiceless sonorants in medial position paired up with their voiced counterparts. As pointed out by Jessen and Pétursson (1998, 43), the palatal and velar nasals only occur before palatal and velar stops respectively and can therefore be regarded as allophonic variations: Apart from the bilabial nasal, [n] and [n̥] occur in all other positions except before palatal and velar stops.

(6)	*hendi*	[hɛntɪ]	(hand)	*henti*	[hɛn̥tɪ]	(to dispose of)
	eldi	[ɛltɪ]	(to cook)	*elti*	[ɛl̥tɪ]	(to chase)
	björg	[pjørk]	(rescue)	*björk*	[pjør̥k]	(birch)
	lamba	[lampa]	(lamb)	*lampi*	[lam̥pɪ]	(lamp)
	banginn	[paʊɲcɪn]	(afraid)	*bankinn*	[paʊɲ̊cɪn]	(bank)
	langa	[laʊŋka]	(to long for)	*banka*	[paʊŋ̊ka]	(to knock)

Voiceless sonorants in post-vocalic position followed by a stop are usually considered as an allophonic variation of the underlying sonorant phoneme (Haugen 1958; Rögnvaldsson 1993; Thráinsson 1978; Hoole 1987, and others). According to this view, the sonorant is devoiced by a shift of the aspiration of the following stop parallel to the preaspiration phenomenon. It has to be noted, however, that the occurrence of voiceless medial sonorants is subject to dialect as shown in (7). In this, the phenomenon of voiceless sonorants or sonorant devoicing is different from that of preaspiration which occurs independently of dialect.

		northern Icel.	southern Icel.
(7)	*henti*	[hɛntʰɪ]	[hɛn̥tɪ]
	elti	[ɛltʰɪ]	[ɛl̥tɪ]
	björk	[pjørkʰ]	[pjør̥k]

Word initial voiceless sonorants [n̥ l̥ r̥] (orthographic: *hn-*, *hl-*, *hr-*) occur in pre-vocalic position in Icelandic where they contrast with with [n l r], see (8).

	nýta	[niːta]	(to use)	*hnýta*	[n̥iːta]	(to knot)
(8)	*líða*	[liːða]	(to pass (time))	*hlíða*	[l̥iːða]	(slope)
	rífa	[riːva]	(to rip)	*hrífa*	[r̥iːva]	(to enchant)

Again, the voiceless sonorants are usually considered to be allophones of the underlying sonorant phoneme. Phonemically, they are represented as sequences /hn hl hr/, (Haugen 1958; Árnason 1986).According to Rögnvaldsson (1993), in the sequence /h/ + sonorant the feature [+SPREAD GLOTTIS] of /h/ is transfered to the following sonorant while /h/ itself is elided.

Jessen and Pétursson (1998), opposing the widely accepted view of an allophonic status of voiceless sonorants, believe that Icelandic voiceless sonorants can be regarded as phonemes. In their study they concentrate on voiceless nasals in Icelandic but explicitly point out that some of the phonetic and phonological aspects apply to other sonorants as well. They mainly highlight the fact that so far phonological analysis largely concentrated on voiceless sonorants in medial position while there seemed "to be only limited awareness of the fact that voiced and voiceless alveolar nasals contrast word-initially in Icelandic (Jessen and Pétursson 1998, 48)". With regard to the medial context, they emphasise that voiceless sonorants and preaspiration are not completely parallel, as shown above.

2.2 Phonation

As part of this study aims to investigate the phonation types of the sonorants in question, a short summary of voice source variation relevant for this study will be given here. Gordon and Ladefoged (2001) describe a simplified continuum by means of the glottal aperture ranging from most closed to most open and phonation ranging from voicelessness via breathy voice, modal voice and creaky voice to the glottal stop respectively. The phonation types relevant for this study range from voicelessness to modal voice, as voiceless sonorants were found to be phonated in a rather lax way sometimes as stated in section 1. Laver (1980) proposes three parameters describing the glottal configuration to specify phonation types: longitudinal tension, medial compression and adductive tension, as displayed in figure 1.

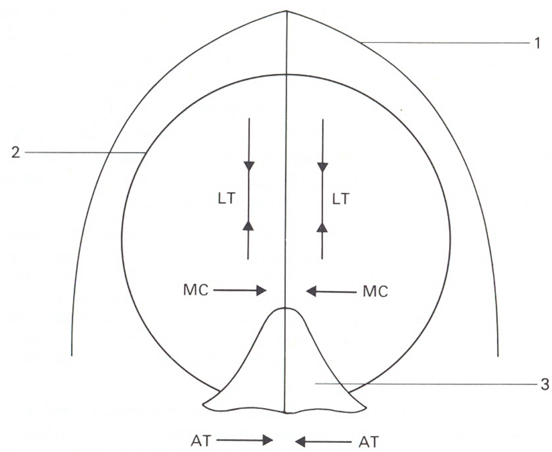


Figure 1: Geometrical relations between three laryngeal parameters (Laver 1980, 109):
LT - longitudinal tension, MC - medial compression, AT - adductive tension.
1 thyroid cartilage, 2 cricoid cartilage, 3 arytenoid cartilages.

Modal voice Among voiced phonation types modal voice is the most common type in the languages of the world (Catford 1977) and has been defined as a neutral, state by (Chomsky and Halle 1968) among others. Accordingly, it serves as the fundament on which the descriptions of other phonation type are built upon. Phonation is usually regarded as cyclical: The glottis is completely closed by medium adductive tension, medium medial compression and medium longitudinal tension. The pulmonic airstream continuously raises until the glottal tension is overpowered causing glottal aperture through which air escapes. The Bernoulli effect, assisted by the sudden reduction of subglottal pressure, causes the vocal-cord to close the glottis, thus closing the cycle. The percentage of glottal aperture within each cycle is called *open quotient* (OQ), commonly used to characterise voice quality (Ni Chasaide 1987).

Breathy voice Minimal adductive tension, weak medial compression and low longitudinal tension characterise breathy voice. During the vocal-cord vibration the glottis is never completely closed, causing considerable air consumption. Accordingly, the OQ is very high for breathy phonation.

Whispered voice / whispery voice Medium longitudinal tension, medium to high medial compression and low adductive tension cause a triangular opening of the cartilaginous glottis. Sufficiently high transglottal airstream causes turbulence and generates the hissing sound, typical for whispered phonation. Lesser medial compression can allow for additional vocal-cord vibration. The outcome is called whispery phonation. Whispery phonation has a lower OQ than breathy phonation.

Voicelessness The main feature of voicelessness is that no vocal-cord vibration is present. For voiceless sounds, except for the glottal stop the glottis is spread, i.e. longitudinal tension, medial compression and adductive tension are minimal. Depending on the glottal area and the transglottal airstream turbulence/frication can arise (breath phonation, Catford 1977).

A way to distinguish between phonation types is to analyse the difference of the first two harmonics (H1-H2) of the spectrum of a sound, as this difference is correlated to the open quotient (Holmberg et al. 1995). According to Ni Chasaide (1987, 442), a “very dominant H1 has been widely found to be highly correlated with a breathy mode of phonation whereas a relatively strong H2 can be correlated with tense or creaky voice.” For rather lax phonation types, the H1-H2 difference can be expected to be less than for modal phonation.

3 Method

3.1 Experimental setup

Two native speakers of Icelandic (f. 26 “theo” / m. 27 “snorri”) were recorded by means of photoglottography (PGG), video (VID) and audio (AU). Both are of southern Icelandic origin and speak the *linmæli* dialect.

Figure 2 shows the setup used in this investigation. For the PGG recordings, an endoscope was introduced through the nasal cavity into the pharyngeal cavity just above the glottis. To monitor the light falling through the glottis, a phototransistor was placed between the thyroid cartilage and the cricoid cartilage (PGG1) and another one just below the cricoid cartilage (PGG2). The signals obtained from the phototransistors were recorded on a digital audio tape

(DAT) recorder along with the audio signal. A camera attached to the ocular of the endoscope was used to monitor the recordings and to record them on a digital video recorder. The strength of the signals detected by light sensors is correlated to the glottal area.

The setup of the recordings imposes restrictions to linguistic material, as the endoscope is prone to displacement by movements of the tongue root in dorsal consonants and back vowels and by velar activity in nasal environments. These restrictions are important for quantitative analysis of the amplitude whereas derived timing relationships are less affected by changes in the distance between the light source and the phototransistors.

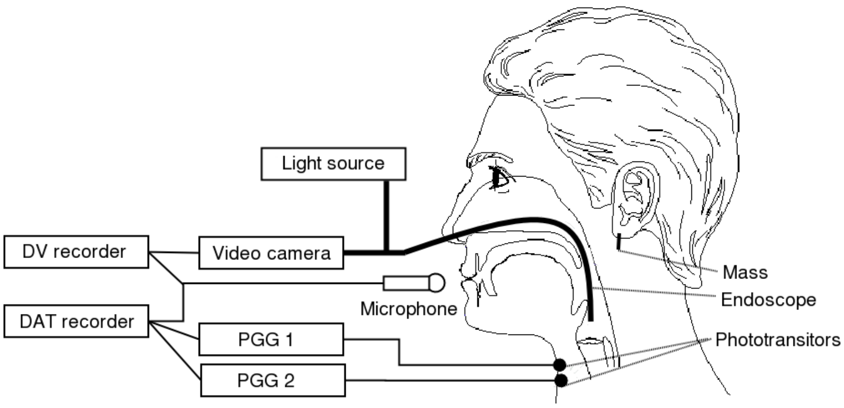


Figure 2: Experimental setup for PGG recordings with two separate phototransistors, audio recordings and video recordings

3.2 Material

The data corpus consists of 22 Icelandic words, some of them inflected to provide minimal pairs and/or to comply with the restrictions imposed by the recording method (see below). All words were embedded into the carrier sentence: "Seppi, lestu ______ fyrir mig." [sɛʰpɪ lɛstʏ ______ fɪrɪr mɪχ] (Seppi, read ______ for me). Table I lists the stimuli relevant for this study. None of the words contain back vowels or dorsal consonants as such sounds would be likely to disturb the PGG recording. In spite of the restrictions mentioned above, nasals were included in the experiment, because they are of main interest in this study. However, there will be no quantitative analysis of PGG amplitudes, so this restriction may be disregarded here.

The first 10 stimuli form pairs displaying the voicing contrast in either the initial or the medial position. For the trill sonorant, no proper stimuli could be found as the contrast only appears in combination with velar stops, see example 1b. The stimuli in the control section (c) are used for comparison only. Two stimuli have voiceless fricatives in initial position, one has a voiceless fricative in medial pre-stop position. As these contexts are identical to those of the sonorants in the initial (a) and medial (b) sections of table I, articulatory behaviour of voiceless sonorants vs. voiceless fricatives can be compared. Furthermore, there are two words with preaspiration in medial position, allowing for comparison with voiceless medial sonorants. The last two stimuli contain a long unaspirated stop and a preaspirated stop respectively in medial position. This pair constitutes a voiced vs. voiceless contrast in the same manner as for the sonorants, see above.

The stimuli were presented to the subjects in ten repetitions in randomised order.

TABLE I: Stimuli used in this study with transcription and translation. Voiced and voiceless sonorants in initial position are shown in the upper part of the table. The middle part shows the medial position. The lower part lists stimuli containing fricatives or preaspiration used for comparison.

	Stimulus	Transcription	Translation
(a) initial	nýta	niːta	to use (inf.)
	hnýta	n̥iːta	to knot (inf.)
	líða	liːða	to feel (inf.)
	hlíða	l̥iːða	slope (gen. pl.)
	rífa	riːva	to rip (inf.)
	hrífa	r̥iːva	to enchant (inf.)
(b) medial	hendi	hɛntɪ	hand (dat. sg.)
	henti	hen̥tɪ	to dispose of (1.sg. past)
	eldi	ɛltɪ	to cook (1.sg. past)
	elti	ɛl̥tɪ	to chase (1.sg. past)
(c) control	fíla	fiːla	to like (inf., colloquial)
	sími	siːmɪ	telephone (nom. sg.)
	lýsti	listɪ	to illuminate (1.sg. past)
	hitti	hi^{h}tɪ	to meet (1.sg. past)
	breytti	prɛɪ^{h}tɪ	to change (1.sg. past)
	breyddi	prɛɪtːɪ	to spread (1.sg. past)

3.3 Data processing

The raw PGG1 and PGG2 data were downsampled to 3 kHz. Additionally, they were filtered by a 60 Hz low pass eliminating the rapid oscillations of voiced intervals to obtain the filtered, smoothed PGG data: PGGF1 and PGGF2. PGGF1 and PGGF2 were then differentiated resulting in velocity signals: PGGV1 and PGGV2.

The zero crossing rate was derived from the audio signal using the tkassp software of the Emu Speech Database System (Cassidy and Harrington 1996). The H1-H2 difference was obtained with the help of the wavesurfer software (Sjolander and Beskow 2005) by using the mean spectrum of a selected segment and measuring the first two harmonics manually. To eliminate possible influences of formants, a correction formula proposed by Iseli and Alwan (2004) was applied to the H1-H2 difference.

Acoustic labelling was done using the xassp software (IPdS Kiel 2005). The label files were then converted to be accessible by the Emu Speech Database System. The following events were labelled in the PGG signals, see figure 3:

`on`: onset of glottal opening: marked by a local minimum in the PGG signal. Coincides with the offset of the preceding glottal opening, therefore inappropriate for marking the onset of the devoicing gesture.

`on10`: 10% threshold of glottal opening: 10% threshold of glottal closing: 10% of the difference of the amplitudes of `on` and `max`. Marks the onset of the devoicing gesture.

The onset of the glottal devoicing gesture has been found to be a robust difference between fricatives and stops, emerging earlier for fricatives than for stops, relative to the formation of oral closure (Hoole 1999).

`xv1`: peak opening velocity: local maximum in the PGGV signal between `on` and `max`.

`max`: peak glottal opening: local maximum in the PGG signal.

Peak glottal opening (PGO) within fricative-stop clusters has been found to occur within the fricative (Hoole et al. 2003), but not necessarily at its center as formulated by Browman and Goldstein 1986, while in sonorant-stop clusters PGO is aligned with the burst of the stop, if a devoicing gesture is present.

`xv2`: peak closing velocity: local maximum in the PGGV signal between `max` and `off`.

`off10`: 10% threshold of glottal closing: 10% of the difference of the amplitudes of `max` and `off`. Marks the offset of the devoicing gesture.

`off`: offset of glottal closure: local minimum in PGG signal.

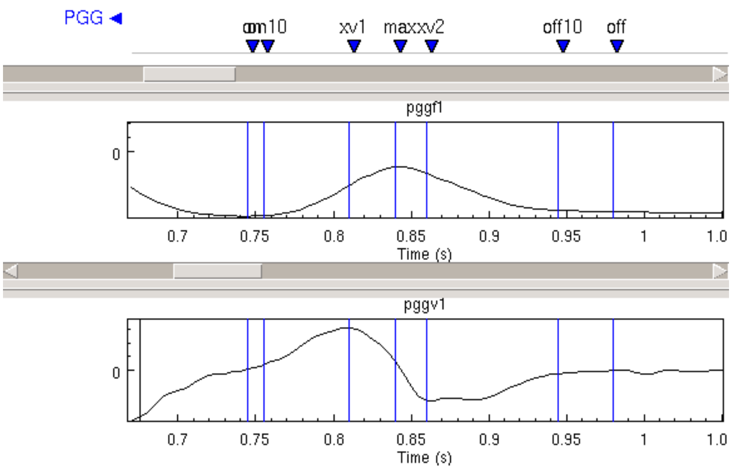


Figure 3: PGG events aligned with the filtered PGG (PGGF) and the velocity signal (PGGV).

To facilitate this task and find the most precise times of these events they were determined automatically using a custom module for Emu. It has to be noted that only the PGG2 signal (sensor below cricoid cartilage) was evaluated here, because it reacts more sensitively to changes of the rear part of the glottis and is thus more suitable for investigating devoicing contexts as pointed out by Löfqvist and Yoshioka (1980, p. 798).

For statistics the R[3] software in combination with the Emu/R package[4] was used. All samples were tested for normal distribution with the Wilks Shapiro test for normality. If the samples were normally distributed they tested with paired two-sample t-tests, else with the Wilcoxon-Mann-Whitney test.

[3] `http://www.r-project.org`
[4] `http://emu.sourceforge.net/emu-splus.shtml`

4 Results

4.1 Duration

The differences between the durations of voiced vs. voiceless sounds were investigated separately for subject, context and sonorant pair (e.g. [n] vs. [n̥]). As all samples were normally distributed, two sample t-tests were applied. The results are shown in table II

Both speakers produced the voiceless nasals in either context with significantly greater duration than voiced nasals. Subject *theo* distinguished [l] and [l̥] in the same manner as she did for the nasals. Subject *snorri*, did not make this distinction in the medial context. In initial position, both subjects produced the voiceless trill with significantly greater duration than the voiced trill. As pointed out in section 3.2, the trills were not investigated in the medial context.

Summary

- Initial [n̥], [l̥] and [r̥] are always longer than their voiced counterparts.
- The same holds for the medial contexts with the exception of one subject's laterals.

TABLE II: Durations of voiced vs. voiceless sonorants in initial an d medial position

		voiced		voiceless				
Context	Subject	n	Mean ms	n	Mean ms	t-test result		
[n] vs. [n̥]								
initial	theo	9	98.7	9	124.5	$t = -8.102$	$df = 8$	$p < 0.001$
	snorri	10	65.3	10	104.9	$t = -5.9901$	$df = 9$	$p < 0.001$
medial	theo	9	78.2	9	109.4	$t = -11.6332$	$df = 8$	$p < 0.001$
	snorri	10	62.9	10	82.5	$t = -4.3641$	$df = 9$	$p < 0.01$
[l] vs. [l̥]								
initial	theo	7	88.0	7	123.7	$t = -7.9417$	$df = 6$	$p < 0.001$
	snorri	10	65.3	10	83.9	$t = -5.1801$	$df = 9$	$p < 0.05$
medial	theo	8	78.9	8	93.5	$t = -2.4652$	$df = 7$	$p < 0.05$
	snorri	9	57.5	9	59.3	$t = 0.7444$	$df = 8$	$\mathbf{p > 0.05}$
[r] vs. [r̥]								
initial	theo	8	50.8	8	89.6	$t = -8.2851$	$df = 7$	$p < 0.001$
	snorri	9	32.5	9	61.4	$t = -7.4728$	$df = 8$	$p < 0.001$

4.2 H1-H2 Difference

To obtain the H1-H2 Difference - a larger value indicating a breathy, a lower value indicating a creaky mode of phonation - the amplitude and frequency of the first two harmonics of initial [n n̥] and [l l̥] were used. These values could be measured without problems to a large extent as the segments provided a sufficient amount of voicing. Only a few tokens of [l̥] constituted a problem due to devoicing. To obtain values for the trills required considerably more effort: H2 values often had to be estimated while H1 was clearly visible. As only very few values were obtainable for the voiceless trills of subject *theo*, they were not evaluated statistically. For statistical analysis, the Wilcoxon-Mann-Whitney test was applied to the H1-H2 differences because the samples did not show normal distribution. For all sonorant pairs, the means of the H1-H2 differences are displayed in table III.

TABLE III: Means of the H1-H2 differences for the sonorant pairs in initial position. Calculated from corrected values (Iseli and Alwan 2004).

	mean H1-H2 (dB)	
	snorri	theo
[n]	-3.08	13.44
[n̥]	6.54	18.58
[l]	-3.95	15.17
[l̥]	6.38	18.17
[r]	1.45	—
[r̥]	11.95	—

For both speakers, the H1-H2 difference of voiceless nasals was significantly higher than of the voiced nasals: *theo*: W = 0, p < 0.001; *snorri*: W = 0, p < 0.001. The H1-H2 differences of the laterals are similar to those of the nasals. For *theo*: W = 9, p < 0.05; for *snorri*: W = 1, p < 0.001. The H1-H2 differences of voiceless trills of the subject *snorri* were significantly higher than their voiced counterparts as well: W = 3, p < 0.001.

Table III shows statistical results of comparisons of H1-H2 differences between the nasals and laterals. As the data were not normallly distributed, the Wilcoxon-Mann-Whitney test was used. No statistical differences could be found.

Summary

- The H1-H2 difference was significantly higher for initial [n̥], [l̥] and [r̥] than for their voiced counterparts, indicating that the voiceless sonorants are phonated in a more lax way than the voiced sonorants.
- There is no statistical difference between H1-H2 differences of laterals and nasals.

4.3 Zero crossing rate

For each relevant segment, the peak zero crossing rate (zcr) - large values indicating frication - was calculated. Figure 4 shows the means of these values separated by sonorant type, context,

TABLE IV: Intra speaker comparison on different sonarants with the same voicing status.

subject	comparison	stat. results
theo	[n] vs. [l]	$W = 25, \quad p > 0,05$
	[n̥] vs. [l̥]	$W = 59, \quad p > 0.05$
snorri	[n] vs. [l]	$W = 49, \quad p > 0.05$
	[n̥] vs. [l̥]	$W = 64, \quad p > 0.05$

subject and phonation together with the fricative [s] for comparison. As these values were not distributed normally, again the Wilcoxon-Mann-Whitney test was used.

Medial voiceless nasals have significantly higher peak zcr than voiced:*theo* $W = 0$, $p < 0.001$; *snorri* $W = 5$, $p < 0.001$. In the initial context, however, there was no significance for *theo* ($W = 23$, $p > 0.5$) while *snorri* has significantly higher peak zcr for voiced nasals ($W = 88$, $p < 0.01$).

Both subjects had significantly higher zcr in voiceless initial laterals than in voiced. *theo*: $W = 0$, $p < 0.001$; *snorri*: $W = 82$, $p < 0.05$. In the medial context, peak zcr was significantly higher for voiceless laterals, too. *theo*: $W = 0$, $p < 0.001$; *snorri*: $W = 0$, $p < 0.001$. Laterals appear to have higher zcr than nasals in general.

The initial voiceless trill had significantly higher zcr than the voiced trill (*theo*: $W = 0$, $p < 0.001$; *snorri*: $W = 2$, $p < 0.001$). Compared with initial nasals and laterals, the trill had the highest zcr.

Summary

- The voiceless nasals, laterals and trill generally show higher zcr than the voiced although never as high as for [s]
- In medial context, zcr of voiceless sonorants is higher than in initial position.
- zcr appears to be dependent of the sonorant type: zcr(n̥) < zcr(l̥) < zcr(r̥)

4.4 PGG

For analysis of the PGG data, only voiceless segments were investigated as the voiced sound do not have a devoicing gesture of the glottis. Additionally, the fricatives [f], [s] and [h] in initial position and [s] and preaspiration (+h) in medial context were included.

To illustrate glottal behaviour as investigated in this study, figures 5 and 6 are included. Figure 5 shows spectrograms with superimposed PGG2 data plots for medial laterals and dental nasals both voiced and voiceless ((a) - (d)). Additionally, (f) and (e) in the same figure show a sequence of short vowel + long stop with an aspiration contrast, parallel to the voiced-voiceless contrast of the sonorants. In the same way figure 6 shows laterals, nasals and trills in initial position.

4.4.1 Timing

As a measure of temporal alignment of the peak glottal opening within a segment the quotient of segment duration and the interval from segment start to peak glottal opening was calculated for each segment. Figure 7 shows the means and standard deviations of this quotient in initial position, figure 8 in medial position, separated by sound and subject.

Peak glottal opening in initial position occurs within the first half of the segment. In medial position, it occurs in the second half or even, for preaspiration, after the segment boundary. However, for the voiceless sonorants in medial position, peak glottal opening is aligned with the sonorant rather than with the subsequent stop, similar to the alignment in fricative-stop clusters.

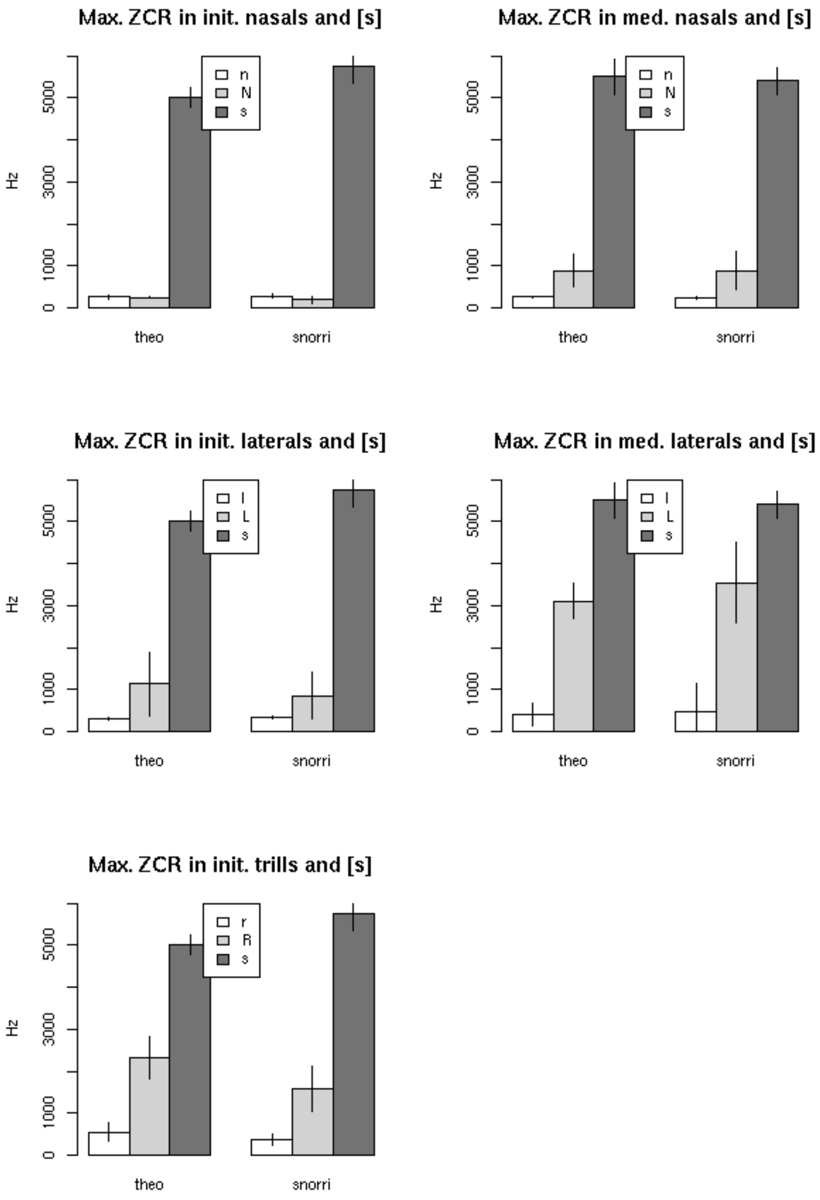
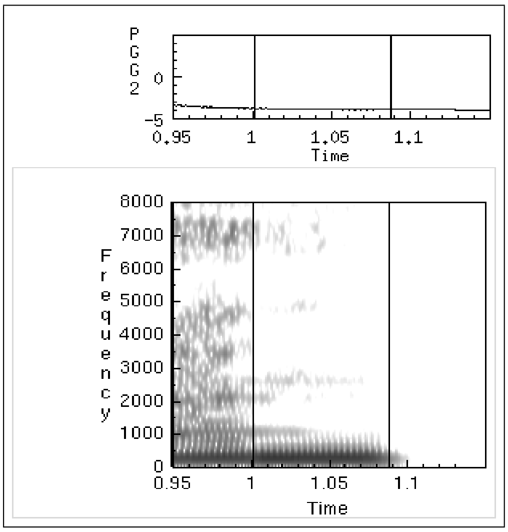
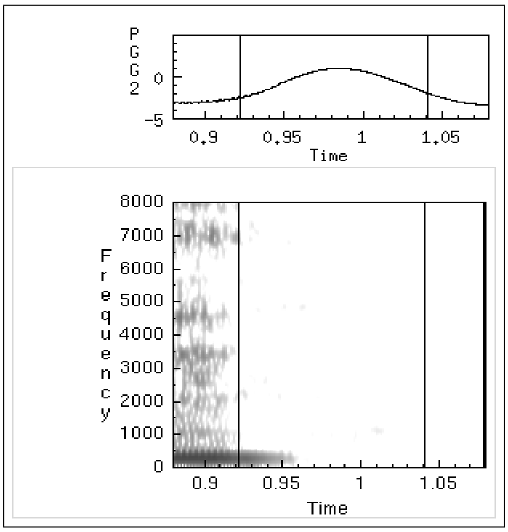


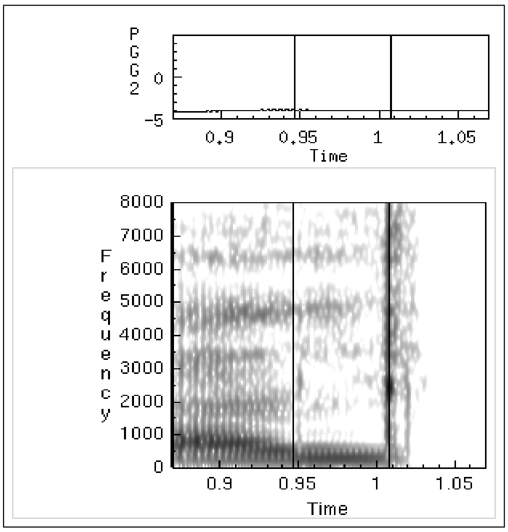
Figure 4: Plots of the means of peak zero crossing rate separated by sonorant type, context, subject and phonation. For comparison, values for [s] in identical contexts are included.



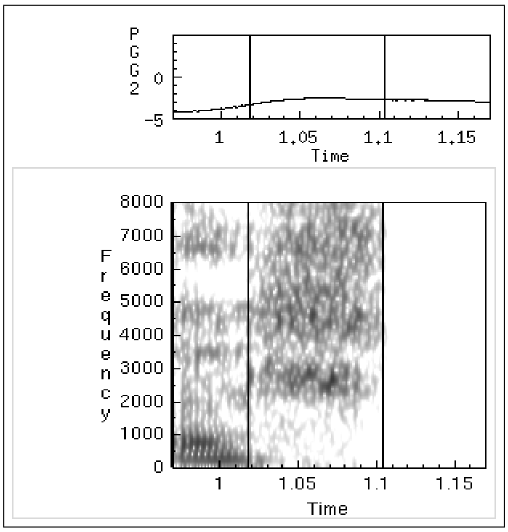
(a) [n]: in *lestu he**n**di* [lɛstʏ hɛ**nt**ɪ]



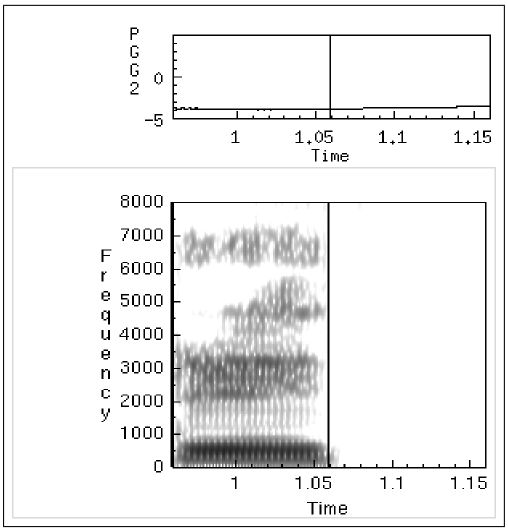
(b) [n̥]: in *lestu he**n**ti* [lɛstʏ hɛ**n̥t**ɪ]



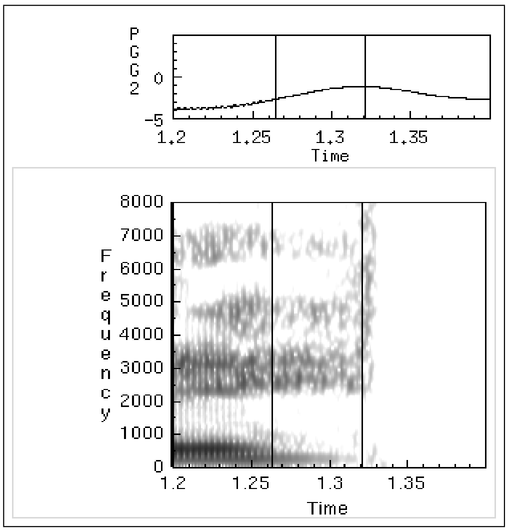
(c) [l]: in *lestu e**l**di* [lɛstʏ ɛ**l**tɪ]



(d) [l̥]: in *lestu e**l**ti* [lɛstʏ ɛ**l̥**tɪ]

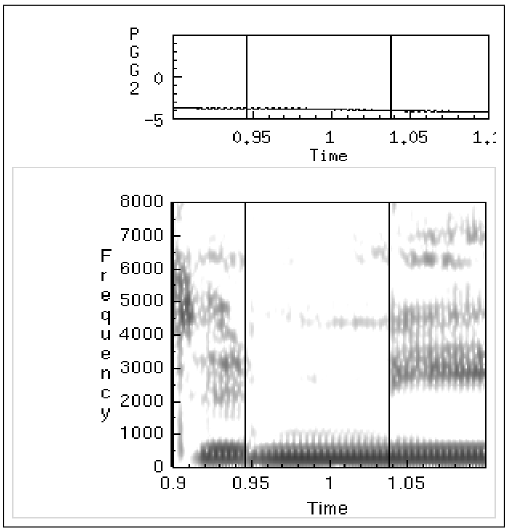


(e) ∅: in *lestu brei**dd**i* [lɛstʏ brɛɪtːɪ]

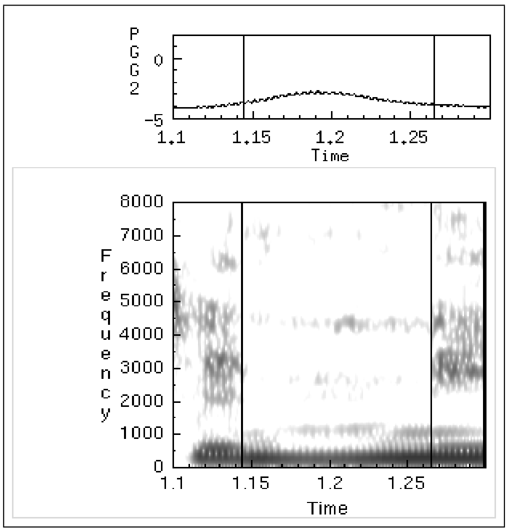


(f) [ʰ]: in *lestu brey**tt**i* [lɛstʏ prɛɪ**ʰ**tɪ]

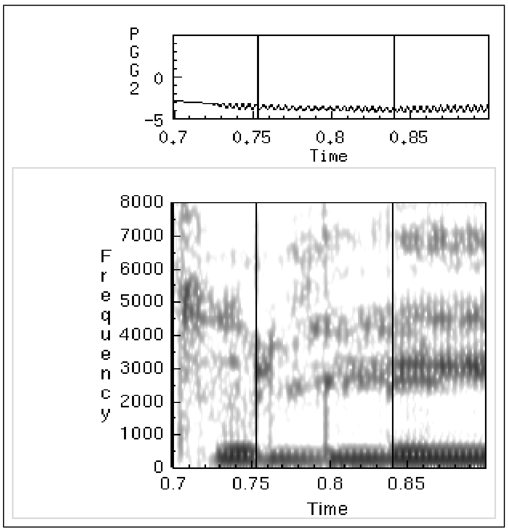
Figure 5: Medial position: plots of PGG2 and spectrogram (time in s, frequency in Hz. (a) through (d) show the contrast of voiced and voiceless sonorants with borders of the segment given in bold letters in the subcaption. (e) shows an unaspirated long lenis stop opposed to a preaspirated long fortis stop in (f).



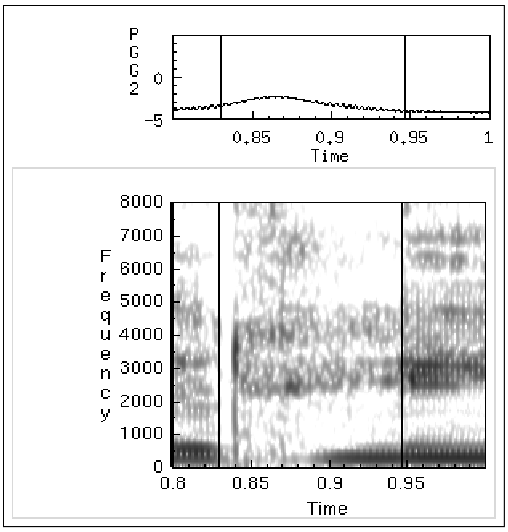
(a) [n] in: *lestu* **n***ýta* [lɛstʏ niːta]



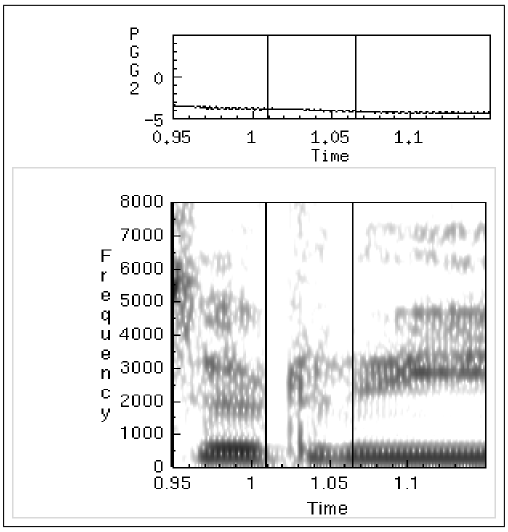
(b) [n̥] in: *lestu* **hn***ýta* [lɛstʏ n̥iːta]



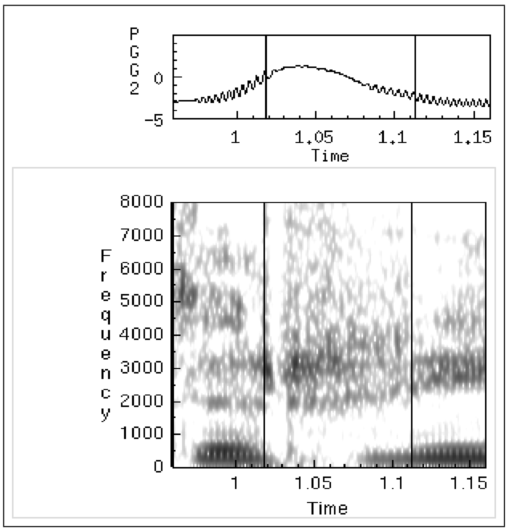
(c) [l] in: *lestu* **l***íða* [lɛstʏ liːða]



(d) [l̥] in: *lestu* **hl***íða* [lɛstʏ l̥iːða]



(e) [r] in: *lestu* **r***ífa* [lɛstʏ riːva]



(f) [r̥] in: *lestu* **hr***ífa* [lɛstʏ r̥iːva]

Figure 6: Initial position: plots of PGG2 and spectrogram (time in s, frequency in Hz. (a) through (f) show the contrast of voiced and voiceless sonorants with borders of the segment given in bold in the subcaption.

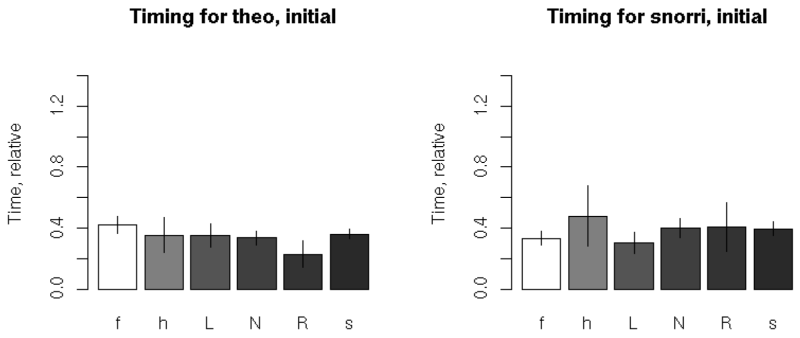


Figure 7: Temporal alignment of peak glottal opening within the underlying segment, initial context, separated by subject.

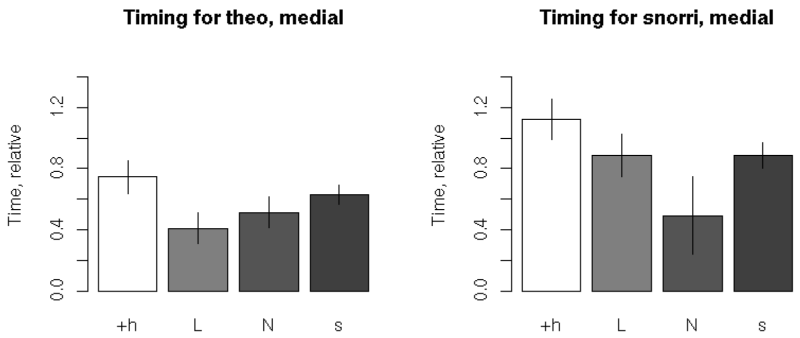


Figure 8: Temporal alignment of peak glottal opening within the underlying segment, medial context, separated by subject.

4.4.2 Properties of the glottal opening phase

For a qualitative description, the (raw) PGG signals for initial [n̥] [l̥] [r̥] [h] and medial [n̥] [l̥] [^{h}] were examined for two features:

1. Is there a glottal devoicing gesture?

2. Is there continuous oscillation of the vocal cords throughout the segment?

A devoicing gesture was found in all cases. Continuous oscillation occurred in some instances of voiceless sonorants and [h] in initial position, but never in medial position and never for fricatives. Vocal fold oscillation was suspended before peak glottal opening for sonorants and in the beginning of the glottal gesture for fricatives.

In the cases where the vocal cords in initial sonorants did not oscillate continuously, offset of oscillation occurred before peak glottal opening and onset before glottal closure. Table V shows the occurrence of vocal-cord vibration for voiceless sonorants and [h]. [n̥] tends to be more likely produced with vocal-cord vibration than [l̥] while [r̥] is least likely.

TABLE V: Frequency of sounds in initial position with continuous oscillation (+) resp. without oscillation of the vocal cords (-).

	theo		*snorri*	
	+	-	+	-
[h]	28	0	28	0
[f]	0	9	0	10
[s]	0	9	0	10
[n̥]	9	0	10	0
[l̥]	4	3	9	1
[r̥]	2	7	8	3

TABLE VI: Frequency of vocal-cord vibration (V), glottal opening (O) and zero crossing rate (F) for initial [h], [n̥], [l̥] and [r̥].

	theo			*snorri*		
	VO	VOF	OF	VO	VOF	OF
[h]	28	0	0	28	0	0
[n̥]	9	0	0	10	0	0
[l̥]	3	1	3	5	5	0
[r̥]	0	3	6	0	8	3
[f]	0	0	9	0	0	10
[s]	0	0	9	0	0	10

4.4.3 Vocal fold vibration vs. zero crossing rate

Sections 4.3 and 4.4.2 show that the voiceless sonorants can be sorted in the same order either by frequency of partial offset of vocal-cord vibration or by peak zero crossing rate. In this light, the devoicing gestures were again examined, but with simultaneous consideration of zcr. Four combinations of frication (zcr, here F), glottal opening (O) and vocal-cord vibration (V) were observed:

VO: continuous vocal-cord vibration and open glottis, zcr < 1 kHz

VOF: continuous vocal-cord vibration and open glottis, zcr > 1 kHz

OF: open glottis, zcr > 1 kHz

O: open glottis

The fricatives [f] and [s] (initially and medially) showed a broad glottal opening and high zcr (OF). Among the sonorants, [h] and preaspiration the combination of features varied, as shown in tables VI and VII. All instances of initial [h] and [n̥] show continuous vibration and glottal opening. Initial [l̥] can have zcr > 1 kHz as an additional feature while vocal-cord vibration may be suspended. [r̥] always shows high zcr while continuous vocal-cord vibration becomes less frequent.

No continuous vocal-cord vibration was found in the medial context for [n̥], [l̥] and preaspiration. The most frequent combination of features was glottal opening with high zcr (OF). Only few instances of preaspiration and [n̥] had merely a glottal opening and no relevant zcr.

TABLE VII: Frequency of glottal opening (V) and zero crossing rate (F) for medial [n̥] and [l̥] and preaspiration. None of the medial items showed glottal vibration.

	theo		*snorri*	
	O	OF	O	OF
[n̥]	2	8	3	7
PRASP	0	9	4	7
[l̥]	0	8	0	10
[s]	0	9	0	10

Summary

- Glottal opening in voiceless sonorants tends to occur earlier in initial context than in medial context.
- While glottal opening is mandatory for all voiceless sonorants regardless of their position, vocal-cord vibration and frication appear to be dependent on position and sonorant type.
- With regard to the properties examined here, some instances of [l̥] and [r̥] are very similar to fricatives especially in medial position.

5 Discussion

The results of this study regarding question 1, i.e. whether there are multiple machanisms contributing to the distinction of voiced vs. voiceless sonorants, can be summarised as follows:

- The duration of voiceless sonorants is greater than of voiced sonorants.
- In initial position, voiceless sonorants have greater H1-H2 difference than their voiced counterparts.
- Voiceless sonorants have more frication (higher zcr) than voiced sonorants.
- In voiceless sonorants, there is a glottal devoicing gesture while there is none in voiced sonorants.
- In many cases, vocal cords vibrate continuously in spite of the glottal aperture.

The differences in duration of voiced and voiceless nasals confirm the findings of e.g. Pétursson (1977) and Hoole (1987). Allthough the actual segment durations in other studies were not reproduced exactly, it appears safe to say that they are speaker dependent as durations vary between both subjects in this study, too. However, all studies display the same tendency. Similar results were found for the other sonorants. Only laterals in medial position for the speaker *snorri* are an exception. Apart from this, duration appears to be a consistent property of the distinction between voiced and voiceless sonorants.

The results of the H1-H2 analysis of initial sonorants (if measurable) suggest that in voiceless sonorants the vocal cords are adducted in a rather lax way when compared to the voiced sonorants, thus indicating that “voiceless” are produced with *breathy* phonation rather than really *voiceless*. However, this only applies where H1 and H2 could actually be measured. As some tokens of initial [l̥] and all instances of initial [r̥] for subject *theo* had to be omitted from this analysis due to the lack of vocal-cord vibration, it would be untenable to propose a contrast of *voiced vs. breathy* instead of *voiced vs. voiceless*.

As an indicator of frication, the zero crossing rate, especially in medial position, is significantly higher for voiceless sonorants than for voiced sonorants. This difference is less prominent in initial position and nonexistent for nasals. This coincides with the assumption that voiceless laterals and trill are more likely to be produced with frication when articulated with large oral constriction (and maybe velar closure) than a nasal regardless of its articulatory setup. Higher zcr in medial voiceless nasals maybe due to a fortified nasal airstream, as observed in Pétursson (1977), causing slight turbulence in the nasal cavities, but there cannot be a nasal constriction constituting frication. Thus zcr in medial voiceless nasals is lower than zcr in voiceless trill, even those initial position.

The observations recorded from the PGG data correspond with the findings of the H1-H2 analysis, as continuous vocal-cord vibration was present only in those cases where H1 and H2 could be measured. However, the always present devoicing gesture of the glottis demands a reconsideration of the phonation type of those voiceless sonorants produced with vocal-cord vibration, as breathy voice has never been described to include a spread glottis. The choice of PGG sensor may yield an explanation: In this study, only the PGG2 signal was evaluated as mentioned in section 3.3 which is more sensitive to changes in the rear part of the glottis. It would be possible for the light to pass through the open rear part of the glottis while only the front part of the glottis oscillates. Accordingly, only the cartilage-glottis would be open and the phonation type could be identified as *whispery voice*.

Halle and Stevens describe another possibility:

> "[By] rotation and displacement of the arytenoid cartilages, the vocal cords can be displaced outward relative to their positions for normal voicing, leaving a large glottal width. If the vocal-cord stiffness is sufficiently large, the combination of wide glottis and stiff glottal walls inhibits vocal-cord vibration. On the other hand, slackening of the glottal walls by reducing the stiffness can lead to a condition in which vocal-cord vibration will occur, even with a relatively wide glottal opening."
> (Halle and Stevens 1971, 201 f.)

In order to clarify this point, a detailed analysis of PGG1 and video data would be helpful. With the data obtained in this study, the observed mode of phonation can only be defined as being some sort of lax.

Based on the results discussed so far, it appears that there are in fact several articulatory mechanisms which help to distinguish between the voiced and voiceless sonorants. However, some exceptions were noted that deviate from the initial idealised summary. It is therefore not possible to answer the first objective in a simple and uniform way as e.g.: "Voiceless sonorants differ from voiced sonorants by lax phonation, greater duration and frication." Instead, the exceptions have to be viewed with regard to objective 2, wether there are sonorant specific mechanisms, that distinguish voiced from voiceless sonorants.

Sections 4.4.2 and 4.4.3 indicate a sonorant specific correlation between vocal-cord vibration and frication in initial sonorants which can be illustrated as follows:

	lax phonation	⟷	voicelessness
(9)	[n̥]	[l̥]	[r̥]
	no frication	⟷	frication

This illustration is based on means and tendencies and does not attempt to define prototypical pronunciations. Also, the exact positions of [n̥ l̥ r̥] are chosen arbitrarily, but their order corresponds to the insights gained here. Furthermore, the actual articulatory realizations, especially of the lateral, are subject to variability. The trill shows less variability than the lateral. Only the nasal is confined to always having lax phonation and no frication. This variability can be explained with regard to perceptibility: While laterals and trills can be produced with frication, nasals depend on another sound source to be audible. Most likely, this explanation interacts with another, articulatory explanation: The oral constriction necessary for frication may inhibit vocal-cord vibration in the laterals and trills due to a raise of supraglottal pressure. This passive devoicing cannot appear for nasals.

This correlation is not present for the medial context, as vocal-cord vibration was never encountered in medial voiceless sonorants. Even so, a similar tendency can be found here: While there is a large amount of frication for the voiceless laterals in all tokens, voiceless nasals can be produced with glottal opening but without any frication. For the absence of frication in nasals the same reasons as mentioned above apply.

The difference between initial and medial voiceless sonorants is, of course, largely due to their contexts. A voiceless segment can be expected to have a larger glottal opening when followed by a stop than when located between vowels. Although PGG data were not analysed quantitatively, plots of the PGG data confirm this assumption. Consequently, there is more space for phonatory variability in the initial context, as the amplitude of glottal abduction may be sufficiently low to allow for vocal-cord vibration.

6 Conclusion

Voicing alterations in Icelandic sonorants were found to be not at all limited to a voiced-voiceless contrast. Instead, duration, voice quality and frication appear to play an important role in the distinction of these sounds. However, concerning the phonation type in initial voiceless sonorants a more detailed analysis of both PGG signals and the video data might be interesting. As to a great extend the results of this study were gained from acoustic analysis, it might be of benefit to acoustically analyse Icelandic sonorants of more speakers in various prosodic contexts. Such an analysis would also be helpfull to derive a more accurate articulatory study.

This work would not have been possible without the enormous help of Susanne Fuchs, Phil Hoole, Klaus Dahlmeyer and Jörg Dreyer during the recording session. Jonathan Harrington, Adrian Simpson and Michel Scheffers provided very valuable comments. Without Christine Mooshammer, this study would still be a mere idea. Thankyou all very much.

References

Árnason, K. (1986). The segmental and suprasegmental status of preaspiration in modern Icelandic. *Nordic Journal of Linguistics 9*, 1–23.

Botma, B. (2004). *Phonological aspects of nasality – an element-based dependency approach*. Ph. D. thesis, Universiteit Amsterdam, Amsterdam.

Browman, C. and L. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook 3*, 219–252.

Cassidy, S. and J. Harrington (1996). EMU: an enhanced hierarchical speech data management system. In P. McCormack and A. Russels (Eds.), *Proceedings of the Sixth Australian International Conference on Speech Science and Technology*, Adelaide, pp. 361–366. ASSTA.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh Univerity Press.

Chomsky, N. and M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Dankovicová, J. (1999). Czech. In *Handbook of the International Phonetic Association*, pp. 70–73. Cambridge: Cambridge University Press.

Gordon, M. and P. Ladefoged (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics 29*, 383–406.

Halle, M. and K. Stevens (1971). A note on laryngeal features. Quaterly Progress Report 101, Research Laboratory of Electronics, M.I.T., Cambridge, Mass.

Haugen, E. (1958). The phonemics of modern Icelandic. *Language 34*, 55–88.

Holmberg, E. B., R. E. Hillmann, J. S. Perkell, P. C. Guiod, and S. L. Goldman (1995). Comparison among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research 38*, 1212–1223.

Hoole, P. (1987). Velar and glottal activity in a speaker of Icelandic. In *Proceedings of the 11th International Congress of Phonetic Sciences*, Volume 3, Tallinn, Estonia, pp. 31–34. Academy of Sciences of the ESSR.

Hoole, P. (1999). Laryngeal Coarticulation. Section A: Coarticulatory investigations of the devoicing gesture. In W. Hardcastle and N. Hewlett (Eds.), *Coarticulation: Theory, Data and Techniques*, pp. 105–121. Cambridge: Cambridge University Press.

Hoole, P., S. Fuchs, and K. Dahlmeier (2003, December). Interarticulator timing in initial consonant clusters. In *Proceedings of the 6th International Seminar on Speech Production*, Sydney.

IPdS Kiel (2005). xassp. `http://www.ipds.uni-kiel.de/forschung/xassp.en.html`.

Iseli, M. and A. Alwan (2004, May). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. In *Proceedings of ICASSP*, Montreal, Canada, pp. 669–672.

Jessen, M. and M. Pétursson (1998). Voiceless nasal phonemes in Icelandic. *Journal of the International Phonetic Association 28*, 43–53.

Johannson, E. T. (2004). The effects of prosodic position on initial voiceless sonorants in Icelandic. In *Proceedings of the 9th Conference of Nordic Prosody*. Manuskript.

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Löfqvist, A. and H. Yoshioka (1980). Laryngeal activity in Icelandic obstruent production. *Haskins SR 63*, 275–292.

Ni Chasaide, A. (1987). Glottal control of aspiration and of voicelessness. In *Proceedings of the 11th International Congress of Phonetic Sciences*, Volume 6, Tallinn, Estonia, pp. 28–31. Academy of Sciences of the ESSR.

Pétursson, M. (1977). Glottographische Untersuchungen der Anlautgruppen hl-, hr-, hn- im modernen Isländischen. In J.-P. Köster (Ed.), *Miszellen 4*, Volume 22 of *Hamburger phonetische Beiträge*, pp. 113–125. Hamburg: Buske.

Rögnvaldsson, E. (1993). *Íslensk hljóðkerfisfræði [Icelandic phonology]*. Reykjavík: Málsvísindastofnun Háskóla Íslands.

Sjolander, K. and J. Beskow (2005). WaveSurfer-an open source speech tool. `http://www.speech.kth.se/wavesurfer/`.

Thráinsson, H. (1978). On the phonology of Icelandic preaspiration. *Nordic Journal of Linguistics 1*, 3–54.

Weigelt, L. F., S. J. Sadoff, and J. D. Miller (1990). Plosive/fricative distinction: The voicelessness. *Journal of the Acoustical Society of America 87*(6), 2729–37.