Acoustic characteristics of Swedish dorsal fricatives

Ryan Shosted

University of Illinois, 4080 FLB, 707 S Mathews Ave; MC-160, Urbana, IL 61801, USA
rshosted@uiuc.edu

The precise articulatory description of the "voiceless dorsopalatal/velar fricative" or "simultaneous" {$\widehat{\int x}$}, usually transcribed as {ɧ} in IPA notation, is somewhat controversial [5]. The controversy arises from highly variable productions among different speakers and the absence of a phonemic contrast with any other dorsovelar fricative in the language. Moreover, the sound is attested only in some dialects of Swedish [8]. The present study attempts to acoustically differentiate anterior and posterior variants of the fricative {ʃ ɧ x} along with the posterior phoneme {h} and the anterior phoneme {ç}, for comparison. A single speaker who controls various regional dialects of Swedish participated in the study. In Experiment 1, the speaker produced VCV nonsense words with balanced V and alternating C. Multitaper spectral estimates were calculated and measured [1, 2, 11]. The results indicate that {ʃ} can be reliably differentiated from {x} and {ɧ} based on center of gravity measures. However, center of gravity differences between {x} and {ɧ} are not revealing, confirming Lindblad's x-ray tracings [8]. The most robust acoustic difference between the two dorsal fricatives is the presence of erratic, high-amplitude, low-frequency transients during {x}. These may be attributed to relatively unpredictable explosions of saliva and ephemeral lingual contact with the uvula or soft palate, more likely for slightly-retracted {x} than for {ɧ}. Based on the current acoustic model of fricative production, it seems unlikely that {ɧ} is produced with a second simultaneous constriction that is of any acoustic relevance. Further experimentation attempts to discover the acoustic correlates that may be exaggerrated when speakers actively distinguish two anterior and two posterior fricatives. The prospectus for a future sociophonetic study is contemplated.

1 Introduction

1.1 Overview

The unique Swedish fricative, symbolized in IPA notation sometimes as [ɧ] and sometimes as [$\widehat{\int x}$], has proven difficult to describe. It has been observed that the allophonic variation of this sound has to do with regional, socioeconomic, and stylistic concerns which entail a sociophonetic analysis beyond the scope of the current paper. Accordingly, the current study will merely elaborate the acoustic realization of this sound, particularly in contrast with that of other fricatives.

The introduction will review various descriptions of how the fricative is pronounced in different regions and among different groups in Sweden. A series of experiments will then be described. These experiments were designed to assess the acoustic characteristics of the two major realizations of the controversial fricative (one posterior and one anterior), which has various orthographic representations in Swedish. The most common written representation—or perhaps simply the most widely recognized by grammarians and lexicographers—is *sj*. This fricative is commonly referred to as *sje-ljudet*, i.e. 'the *sje* sound'. It contrasts phonemically with another fricative, or so-called *tje-ljudet*. During the experiments, a single speaker of Swedish pronounced words containing the fricatives under investigation. Tapers from the discrete prolate spheroidal series were applied to the excised fricatives and various aspects of the resultant spectrum were measured [1, 2, 11]. The aim is to discover those acoustic correlates that successfully differentiate the various articulations of swedish *sj* from the other fricatives. It is hoped that this study will pave the way for more comprehensive research that will help explain the reason for the fricative's high variability among Swedish dialects.

1.2 Overview of Swedish fricatives

A standard Swedish pronouncing dictionary [10] describes the articulation of the language's fricatives; the relevant ones are discussed here. [ç] (*tje-ljudet*) is described as having a lamino-postalveolar place of articulation, with the tongue blade positioned behind the alveolar ridge. The tongue body rolls forward and approximates the hard palate, resulting in some palatalization. [ʂ̥] is described as an apico- or laminopostalveolar fricative with the tongue tip placed against the back of the alveolar ridge (the transcription of this sound as a retroflex is unaccounted for in [10] and appears elsewhere simply as [ʃ]). [h] is described as a glottal fricative that is often "colored" by surrounding vocalic elements.

Finally, [ɧ] is described as a dorsovelar fricative, i.e. the tongue dorsum is raised very near the soft palate. It is observed that there are several types (*flera slags*) of [ɧ] in Swedish, which are rendered colloquially as *mörka* 'dark' and *ljusa* 'light'.[1] The *ljusa* variant is sometimes indiscernible from anterior [ʂ̥]/[ʃ] and the *mörka* variant is a more posterior fricative, reminiscent of [x]. This leads us to consider whether *mörka* [ɧ] is in fact different from [x] and whether *ljusa* [ɧ] is different from [ʂ̥] or [ʃ].

In the southern city of Lund, the pronunciation of *sje-ljudet* varies between [ɧ] and [ʂ̥] depending on gender and level of education (women with a high degree of education use [ʂ̥] almost exclusively) [4]. Swedish speakers from Jönköping, in the south-central region, vary their pronunciation of the fricative, though some speakers find the posterior articulation to be "vulgar" [9]. In Göteborg, on the western coast, the 'dark' *sj* is said to be produced with the tip of the tongue at the back of the lower teeth, with the tongue body raised and grooved in order to produce a 'heavy hissing' sound as air flows through the long, broad channel. The 'light' (alveolar) *sj* is said to have a considerable degree of lip-rounding. Thus, when lip-rounding is reduced and/or the point of constriction becomes more posterior, the fricative is quite similar to what Holmberg calls *rent velart* 'the

[1] The terms *tjockt* 'thick' and *tunt* 'thin' are also employed [3].

simple velar' of German, viz. [x] [3]. In the region of Göteborg, men have a tendency to produce the velar or *mörka* fricative whereas women are much more likely to produce an alveolar or *ljusa* fricative. In Luleå (northern Sweden), men produce both [ɧ] and [ʂ̜] but women produce only [ʂ̜]. In Malmberget, near the Arctic Circle, only [ɧ] is heard in men's speech, whereas both [ʂ̜] and [ɧ] are heard among women [8]. Throughout Sweden it appears that women only produce the dorsovelar fricative if men do, as well.

Lindblad's [8] is the most thorough articulatory and acoustic description of Swedish fricatives. He accounts for regional and gender-based variation in the production of /ʃ/ and /ç/, i.e. phonemic *sje-ljudet* and *tje-ljudet*. Lindblad cites two different posterior allophones of *sje-ljudet*, viz. [ɧ] and a labialized variant [ɧʷ].[2] He describes the former as having a large jaw angle and unrounded lips, while the latter has a narrow mouth opening, small jaw angle, and close lip rounding. Lindblad argues that the sound source is labiodental for [ɧʷ] and dorsovelar for [ɧ]. Furthermore, he makes the provocative claim that some intermediate variants of the fricative have simultaneous anterior and posterior sound sources.[3] Acoustic models of speech production assume that for fricatives, which are typically produced with a miniscule constriction in the vocal tract, acoustic coupling between the two sides of the constriction is negligible [12]. As a result, the size and shape of the posterior cavity are probably irrelevant to the acoustics of the sound. While in articulatory terms the double constriction may indeed be possible, our present acoustic models of fricative production make Lindblad's multiple-source analysis problematic. It is hard to imagine what the acoustic—and hence perceptual—outcome of a multiply-articulated fricative might be.

2 Methods

The experiments described in this section were designed to differentiate the various posterior and anterior fricatives of Swedish: [h] vs. [x] vs. 'dark' *sje-ljudet* (symbolized henceforth as [ɧ]) and [ç] (*tje-ljudet*) vs. 'light' *sje-ljudet* (symbolized henceforth as [ʃ]). In Experiment 1, VCV nonce tokens (e.g. [iɧi]) were produced without any special emphasis. In Experiment 2, nonce CV syllables were produced under controlled conditions of alternating emphasis. A female speaker of Stockholm Swedish participated in both experiments. Having lived in a northern dialect region of Sweden in her youth, the speaker felt comfortable producing various realizations of *sje-ljudet*, which she was called upon to do in both experiments.

2.1 Experiment 1

The speaker was recorded in a sound-attenuated booth using a Marantz PMD570 professional installation solid-state recorder at a sampling rate of 22.05 kHz. The speaker wore an AKG C520 head-mounted condenser microphone. The microphone, covered with a foam windscreen, was placed approximately 5 cm from the corner of the lips. Signals were pre-amplified using a Grace Design high fidelity pre-amp (model 101). Tokens were nonsense CVC syllables uttered in the carrier phrase, *Jag säger ___ igen* 'I say ___ again'. Tokens included three fricative allophones of *sje-ljudet*, [ʃ ɧ x]. The corpus consisted of 27 tokens.

2.2 Experiment 2

Recording equipment and conditions were as above, except for a samping rate of 44.1 kHz. Tokens in this experiment were balanced sets of CV syllables. In the first part of the experiment, the subject pronounced the syllables in a carrier phrase, *Jag sa ___ igen* 'I said ___ again'. Tokens uttered in this context are said to have 'neutral' emphasis in the analysis. In the second part of the experiment, the subject was asked to make *emphatic contrasts* between the various fricatives, using the carrier sentence, *Jag sa ___ , inte ___* ' I said ___ , not ___'. The subject was instructed to emphasize the first (non-negated) 'word' of each pair. Tokens included two allophones of *sje-ljudet* [ʃ ɧ] (anterior and posterior) along with two other phonemes unrelated to *sje-ljudet*, anterior [ç] and posterior [h]. The nonce pairs are show in Table 1. The corpus consisted of 144 tokens.

Emphatic	Non-emphatic
çi	ʃi
ça	ʃa
çu	ʃu
ʃi	çi
ʃa	ça
ʃu	çu
hi	ɧi
ha	ɧa
hu	ɧu
ɧi	hi
ɧa	ha
ɧu	hu

Table 1: Stimuli for Experiment 2

2.3 Measurements

Multitaper spectral estimates were computed using a sequence of orthogonal tapers specified from discrete prolate spheroidal sequences [1, 2, 11]. This method was chosen to account for the stochastic quality of fricative noise, which gives rise to high variance in spectral shape from one window to the next. The amplitudes of the spectra were normalized such that the maximum amplitude corresponded to a value of 10,000 (this served merely to scale slope values to a manageable number of decimal places).

Least-squares regression lines were fitted to the data below the maximum amplitude and above the maximum amplitude [6]. The resulting values, called LoSlope and HiSlope, are a measurement in dB/Hz that represent

[2] Lindblad uses a symbol somewhat like [ɷ] for this sound.

[3] It is presumably because of this claim that [ʃ͡x] and [ɧ] are equated with one another in the IPA table of consonants.

the change in normalized spectral amplitude over a given frequency bandwidth. LoSlope values should be positive for so-called 'strident' fricatives like [ʃ] and [ç], which have a prominent spectral peak, because the spectral envelope is rising in the low frequencies. LoSlope values should be closer to zero for non-strident fricatives like [x] and [f] where a prominent spectral peak is not likely. A deeply negative value in HiSlope indicates a fast energy drop, indicative of a spectral peak—hence, indicative of a strident fricative. A HiSlope value closer to zero indicates a non-strident fricative.

Multitaper spectral estimates, along with regression lines representing HiSlope and LoSlope are shown in Figures 1–4. In these figures, the fricatives have been drawn from the emphatic stimuli of Experiment 2.

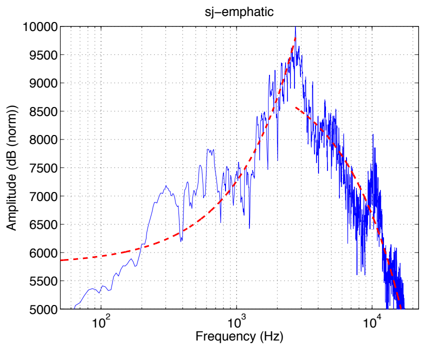

Figure 1: Spectrum of emphatic [ʃ], plotted on a logarithmic axis with normalized amplitude.

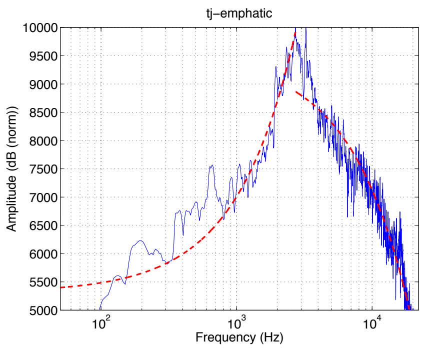

Figure 2: Spectrum of emphatic [ç]. The spectral differences between [ç] and [ʃ] (Figure 1) are not immediately obvious.

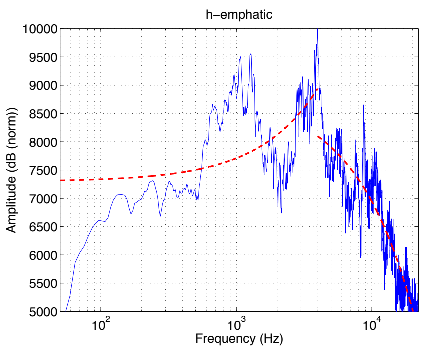

Figure 3: Spectrum of emphatic [h].

Center of gravity measures of the fricative spectra were also calculated. A higher center of gravity is expected

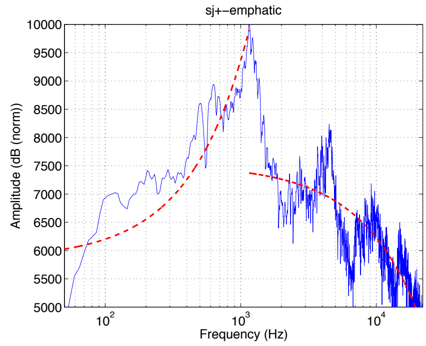

Figure 4: Spectrum of emphatic [ɧ]. The fricative seems more peaked than [h] (Figure 3).

for fricatives with a shorter anterior resonating cavity (hence, the center of gravity for [ʃ] should be considerably greater than the center of gravity for [ɧ]).

3 Results

3.1 Experiment 1

One-way ANOVA showed that center of gravity (COG) was a significant predictor of place of articulation for the various allophonic realizations of *sje-ljudet*, $F(2, 24) = 5.89, p < 0.01$. As shown in Figure 5, the center of gravity is highest for [ʃ], then lower but with much higher variance for [ɧ] and [x]. The post-hoc TukeyHSD test revealed a significant difference for [ʃ] vs. [ɧ] ($p < 0.05$) and [ʃ] vs [x] ($p < 0.05$) but not for [ɧ] vs. [x].

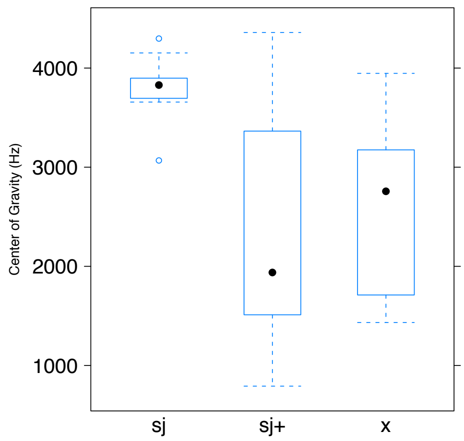

Figure 5: Experiment 1: COG for three different realizations of *sje-ljudet*; [sj]=[ʃ], [sj+]=[ɧ].

HiSlope and LoSlope did not emerge as significant predictors of place of articulation, $F(2, 23) = 0.44, p > 0.05$ and $F(2, 22) = 0.14, p > 0.05$.

3.2 Experiment 2

Neutral context One-way ANOVA showed that center of gravity (COG) was a significant predictor of place of articulation for the Swedish fricatives when CV tokens were uttered without any particular emphasis (neutral context), $F(3,44) = 13.11, p < 0.001$. A TukeyHSD test revealed that there were significant differences ($p < 0.05$) for [ʃ] vs. [h], [ç] vs. [h], and [ɧ] vs. [ʃ]. There were no significant differences in COG for [ʃ] vs. [ç], [ɧ] vs. [ʃ], or [ɧ] vs. [h]. The boxplots for COG are given in Figure 6.

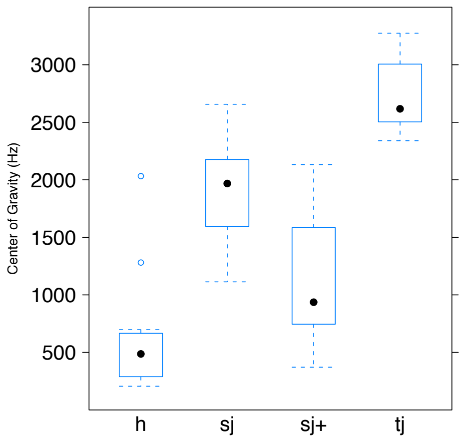

Figure 6: Experiment 2: COG for fricatives in neutral context; [sj]=[ʃ], [sj+]=[ɧ], [tj]=[ç].

HiSlope was found to be a significant predictor of place of articulation, as well, $F(3,44) = 25.04, p < 0.001$. TukeyHSD revealed significant differences between all pairs of fricatives except [ɧ] vs. [h] and [ç] vs. [ʃ]. The boxplots for HiSlope are given in Figure 7.

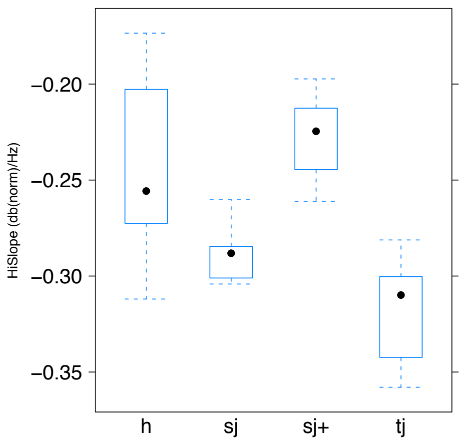

Figure 7: Experiment 2: HiSlope for fricatives in neutral context.

LoSlope was not found to be a significant predictor of place of articulation, $F(3,44) = 1.47, p > 0.05$. TukeyHSD revealed no significant differences for any fricative pairs.

Emphatic vs. unemphatic context ANOVA showed that COG was a significant predictor of both place of articulation, $F(3,92) = 119.34, p < 0.001$, and emphasis condition $F(1,91) = 5.23, p < 0.05$. TukeyHSD revealed significant differences between all pairs except [ɧ] vs. [h] ($p > 0.05$). The unemphatic vs. emphatic pairing had a significant difference of $p < 0.05$. The boxplots are shown in Figure 8.

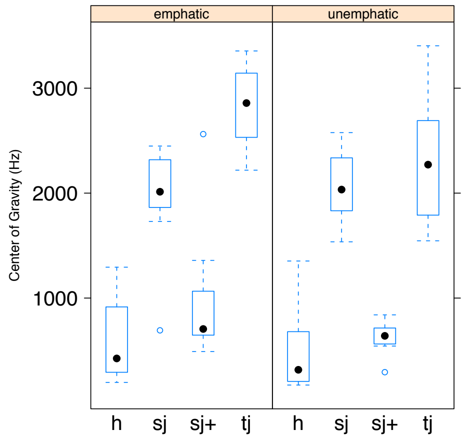

Figure 8: COG of emphatic and unemphatic fricatives.

Hislope was shown to be a significant predictor of fricative, $F(3,92) = 68.84, p < 0.001$, but not emphasis condition $F(1,91) = 0.04, p > 0.05$. Significant differences were detected between all fricative pairs except [ç] vs. [ʃ] and emphatic vs. unemphatic. Boxplots are given in Figure 9.

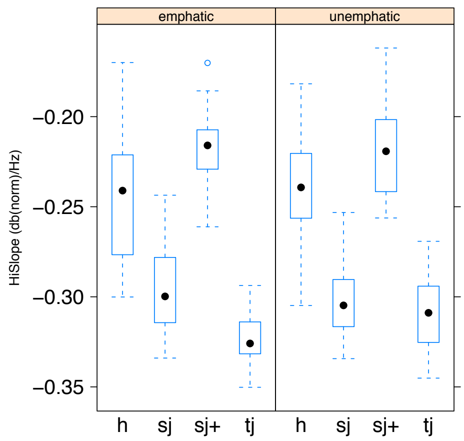

Figure 9: HiSlope of emphatic and unemphatic fricatives.

LoSlope was not found to be a significant predictor of either fricative, $F(3,89) = 2.52, p > 0.05$, or emphasis condition, $F(1,88) = 1.32, p > 0.05$. TukeyHSD returned no significant pairings.

4 Discussion

The results of Experiment 1 did not strongly confirm the role of HiSlope in differentiating the realizations of Swedish *sje-ljudet*. However, this may be due to the fact that there were so few tokens in the experiment. The neutral condition in Experiment 2, which included about twice as many tokens as Experiment 1, was able to strongly confirm the role of center of gravity and HiSlope in differentiating fricatives of Swedish. However, even center of gravity, which was significant in differentiating fricatives in Experiment 1, could not differentiate the *mörka* or 'dark' variant of *sje-ljudet* from the velar fricative [x]. This indicates that the two sounds do not differ greatly from one another and further confirms Lindblad's x-ray tracings of tongue shape, which showed little difference between the two articulations [8, 7]. Qualitative analysis of the sound pressure signal, however, reveals an interesting difference: erratic, high-amplitude, low-frequency transients are commonly seen in recordings of [x]. It has been suggested that these may be indicative of closer contact between the tongue dorsum and soft palate, resulting in minute explosions of saliva or even epiphenomenal contact between tongue and velum or tongue and uvula [8]. This analysis seems consistent with the present data.

The results of Experiment 2 with regard to place of articulation (in the neutral context) were predictable for the posterior fricatives, but not the anterior ones: [ɧ] has a higher center of gravity (hence more anterior place of articulation) than the glottal fricative [h] (which exhibits a relatively large degree of variance). However, [ç] could not be differentiated from [ʃ] using center of gravity, indicating that place of articulation for the two fricatives (i.e. length of the anterior resonating cavity) is not all that different. The results of Experiment 2 (still in the neutral context) suggest that the high-frequency energy drop (HiSLope) may be steeper for [h] than for [ɧ] but not steeper for [ç] than for [ʃ]. This measure also indicates that the posterior fricatives [ɧ h] have a significantly flatter spectrum than their anterior counterparts.

The results of Experiment 2 indicate that center of gravity can be manipulated for purposes of emphasis in fricatives. Center of gravity tended to rise slightly for most fricatives when placed under emphasis but the effect was most noticeable for [ç], i.e. *tje-ljudet*. In the experiment, this fricative was contrasted with [ʃ], i.e. the *ljusa* or 'light' variant of *sje-ljudet*. Thus, it seems reasonably clear that the speaker adopted a strategy of fronting the fricative [ç] to differentiate it from the more retracted [ʃ]. Similar claims cannot be made about the posterior fricatives. Notably, the speaker appears to have adopted no coherent strategy for emphasizing the posterior [ɧ] or 'dark' *sje-ljudet* in contrast to glottal [h].

5 Conclusions

The goal of this study was to establish some acoustic criteria for differentiating the fricatives of Swedish, and the allophonic realizations of *sje-ljudet* in particular. When *sje-ljudet* is realized as an anterior fricative, it is not always differentiable from the other anterior fricative, [ç]. For those speakers of Swedish who do not produce posterior [ɧ] (e.g. women in Göteborg), this means that either the place of articulation for [ç] is under pressure to move further forward in order to accommodate [ʃ] or that the distinction between the two fricatives is collapsing. Further acoustic investigation of this topic may show the activity of a sound change in progress.

Because there are not two phonemic, *posterior* buccal fricatives in Swedish, there is little reason to differentiate [ɧ] from [x] in everyday speech, only perhaps when drawing a distinction between the German velar fricative and the Swedish sound. This hyper-differentiation—the somewhat unsuccessful object of elicitation in Experiment 1—may be the reason for the ephemeral dorsovelar contact and explosions of saliva associated with [x], the variant of *sje-ljudet* attested in Göteborg.

It is interesting to note that the attitude towards 'dark' [ɧ] is negative in at least some regions of Sweden, where it is treated as a "vulgar" speech sound. Throughout the country it is most often associated with the speech of men instead of women. Nevertheless, the spectral qualities of [ɧ] do not readily distinguish it from [x], which may explain why the two realizations vary freely among male speakers in Göteborg. Linguistic attitudes towards the dorsovelar fricative may yet prove an interesting field of sociophonetic study, revealing why speakers opt to modify their pronunciation and how they accommodate existing phonemes when alterations in their phonetic inventory become desirable.

References

[1] O. S. Blacklock, "Characteristics of variation in production of normal and disordered fricatives, using reduced-variance spectral methods", Ph.D. Thesis, University of Southampton (2004)

[2] O. S. Blacklock, C. H. Shadle, "Spectral moments and alternative methods of characterizing fricatives", *J. Acoust. Soc. Am. 113*, 2199 (2003)

[3] B. Holmberg, *Språket i Göteborg*, Lund: Berlingska Boktryckeriet (1976)

[4] I. Ingers, *Språket i Lund: Dialekt och riksspråk i Lund och omnejd*, Stockholm: Norstedt & Söner (1970)

[5] International Phonetic Association, *Handbook of the International Phonetic Association*, Cambridge: Cambridge University Press (1999)

[6] L. M. T. de Jesus, C. H. Shadle, "A parametric study of the spectral characteristics of European Portuguese fricatives", *J. Phonetics 30*, 437–464 (2002)

[7] P. Ladefoged, I. Maddieson, *The Sounds of the World's Languages*, Oxford: Blackwell (1996)

[8] P. Lindblad, "Svenskans sje- och tje-ljud i ett allmänfonetisk perspektiv", *Travaux de l'Institut de Linguistique de Lund 16*, Lund: C. W. K. Gleerup (1980)

[9] E. Lönnerholm, *Språket i Jönköping*, Lund: Berlingska Boktryckeriet (1972)

[10] H. Nygren, ed., *Svenska språknämndens uttalsordbok*, Gjøvik, Norway: Norstedts Ordbok

[11] D. Percival, A. Walden, *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*, Cambridge: Cambridge University Press (1998)

[12] K. N. Stevens, *Acoustic Phonetics*, Cambridge: MIT Press (2000)