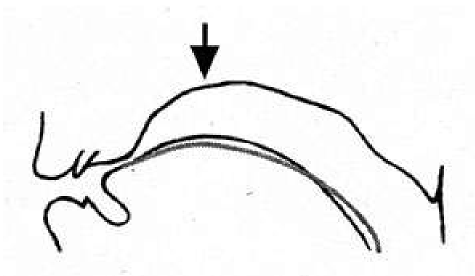
Acoustic analysis of Polish fricatives

Armin Schmidt <armin@ikiro.de>

Torsten Marek <shlomme@gmx.net>

1. Pronounciation of ʃ and ɕ

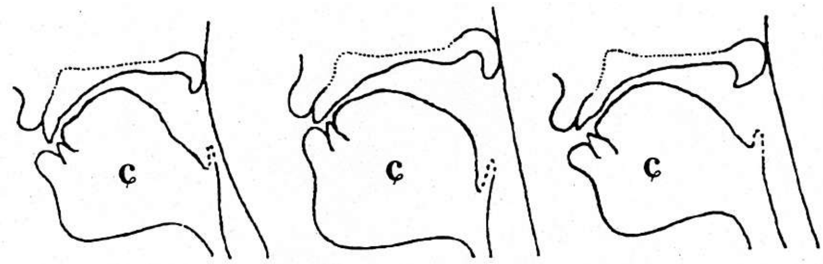
[ʃ] is a palato-alveolar fricative. It is produced with the teeth close together, making it a strident fricative. The tongue creates a constriction in the post-alveolar region whereas the part of the tongue immediately behind this constriction is raised (also referred to as domed). This doming is similar to a small amount of palatalization which is explained in the paragraph on the articulation of [ɕ].



[Figure 1]

Fact but difficult to see in figure 1 is that the front of the tongue is raised with the center being above the level of the sides.

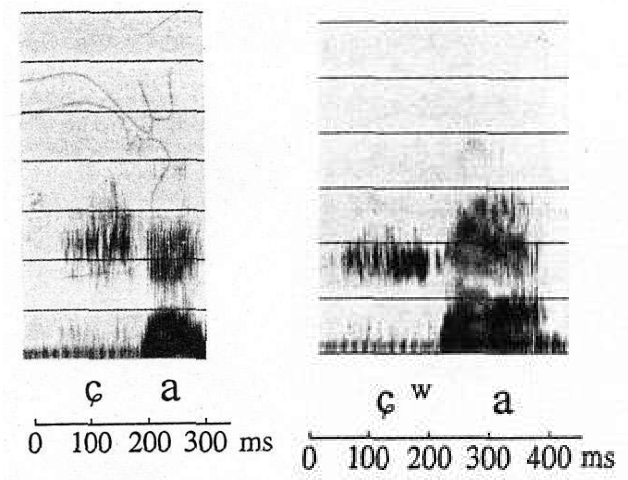
[ɕ] is a palatalized post-alveolar fricative . For its production the upper and lower teeth are fairly close together, for that it is a sibilant fricative. The main source of acoustic energy is the turbulence that arises when the air passes through the nearly clenched teeth.The flat blade of the tongue is raised toward a position that is similar to that of [i]. The lips are spread and the air escapes through the narrow channel between the post-alveolar region, directing the jet of air downward to strike the teeth and thus produce a sibilant sound.



[Figure 2] Tracings from x-rays of different speakers of Mandarin Chinese, articulating [ɕ]. (According to Ladefoged and Maddieson (1996), the production is quite the same as in Polish.)

Blade and body of the tongue are higher than in [ʃ] forming a long and flat constriction. There is contact between the sides of the tongue and the molar teeth so that it could be that, as suggested by Shadle (1985), air turbulences are formed along this wall.

As can be seen in the sample spectrograms below, [ɕ] has its main energy at about 2500 Hz when unrounded and clearly below 2000 when rounded.
[ʃ] differs from [ɕ] in that it has a broader frequency range.



2. The sample words

Sound	Word (Context)	Transcription	Translation
ʃ(sz)	mniszek (i_e)	[mniʃɛk]	herbal remedy
	straszny (a_n)	[straʃnɨ]	horrible
ɕ (ś, si-)	misiek (i_e)	[miɕɛk]	toy
	kwaśne (a_n)	[kvaɕnɛ]	sour

3. Analysis

Spectrogram analysis and formant structure

3.1. To distinguish between the sounds, we examine the formant structure. The values are in Hertz and computed by praat's "Get formant" function (Editor window, formant menu). Due to the facts that the formant structure is very chaotic at the beginning of the fricative, the values are taken from the middle of the second half.
All the spectrograms with formants added are appended.

		Context			
Sound	Formant	i_e		a_n	
ʃ(sz)	F2	1900	2000	1600	1800
	F3	3200	3400	3200	3200
	F4	3900	4000	3900	3900
ɕ (ś, si-)	F2	2600	2700	2400	2500
	F3	3400	3400	3200	3200
	F4	4300	4300	4100	4000

3.2.Similarities

As expected for fricatives, both consonants have strong unstructured noise over (but not always beginning at) 2500 Hz. The third formant of both sounds appears to be around 3.3 kHz, but the formant structure is not always clear enough to make a proper statement.

To compare the length, we measured the time span of the strong noise in all recordings. It lies, except for kwasne (1), which we liberally considered to be an outlier, consistently at 130 ms (the seventh part of a second).

3.3. Differences

Only by looking at the spectrograms, it can be seen that[ɕ] is much stronger between 2 and 3 kHz than [ʃ].

A comparison of the formats shows, that the average F2 of [ʃ] lies approx. at 1.8 kHz (it varies with the context), while F2 of [ɕ] is at 2.6 kHz. Moreover, F2 of [ɕ] is much stronger, which corresponds to the observations made first.

The difference is not that clear for F4, while [ʃ] is tied to 3.9 – 4.0 kHz in all cases, [ɕ] is only around 300 Hz higher (but not in all contexts).

3.4. Variation

Although the second formant is difficult to be spotted in the recordings with [ ], the formant measurement shows that the height of F2 depends on the preceding vowel. If the vowel is low (a), F2 is at ~1.7 kHz for [ʃ] and ~2.45 kHz for [ɕ], after a high vowel the second formant is around 300 Hz higher for both fricatives.

3.5. Comparison with Ladefoged & Maddieson

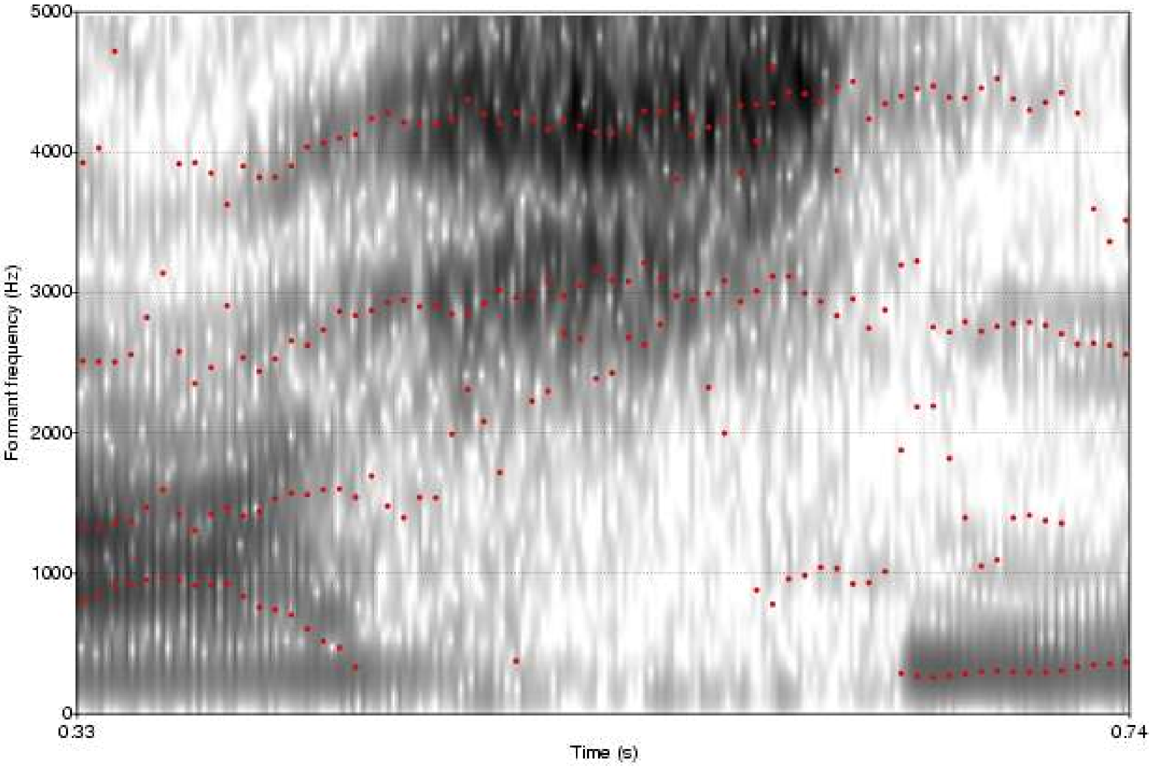
What has been found in [LaMa] and described in 1 mostly corresponds to our own evaluations and recordings. One interesting aspect might be the observation that the main energy of the second formant in our example [kwaɕne] is in a range between 2600 and 2700. This shows not only that Polish [ɕ] is unrounded but might be interpreted in favor of the fact that our native speaker not only not rounds her lips but even pulls back the corners of her mouth towards her ears during production.

3.6. Articulatory aspects

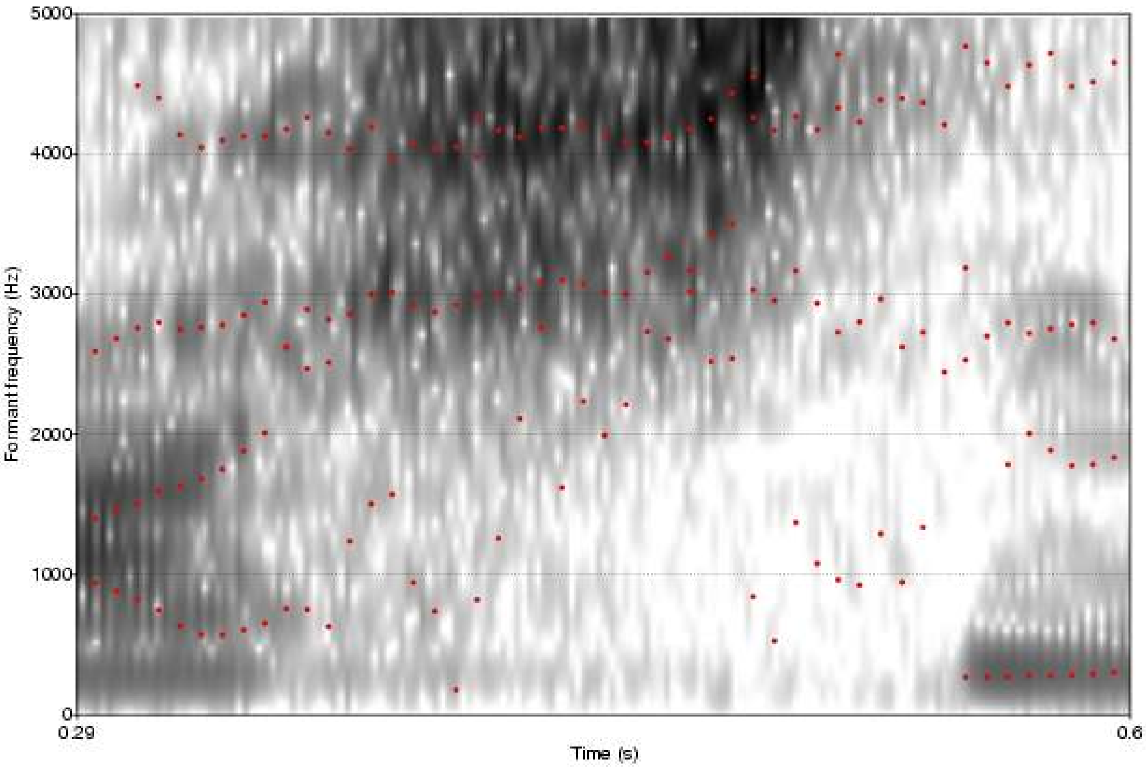
The first aspect is the voicelessness of both sounds, which can be seen by the complete lack of vertical striation and the random patterns of the unstructured noise.
The different frequency of the second formant can be explained by tongue position. During the pronounciation of [ʃ], the tongue forms a larger space in the vocal tract than during the production of [ɕ]. The frequency, then, is expectedly lower in conclusion.

Literature:

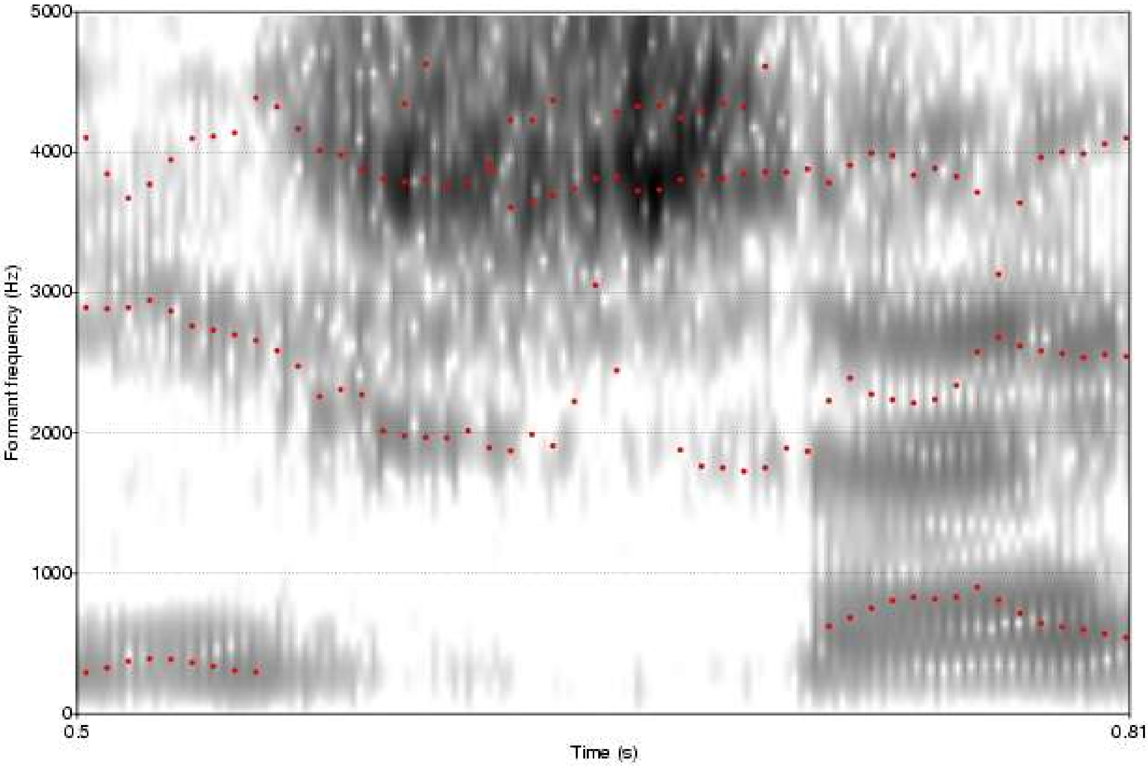
[LaMa]: Ladefoged and Maddieson (1996) "The sounds of the languages of the world"

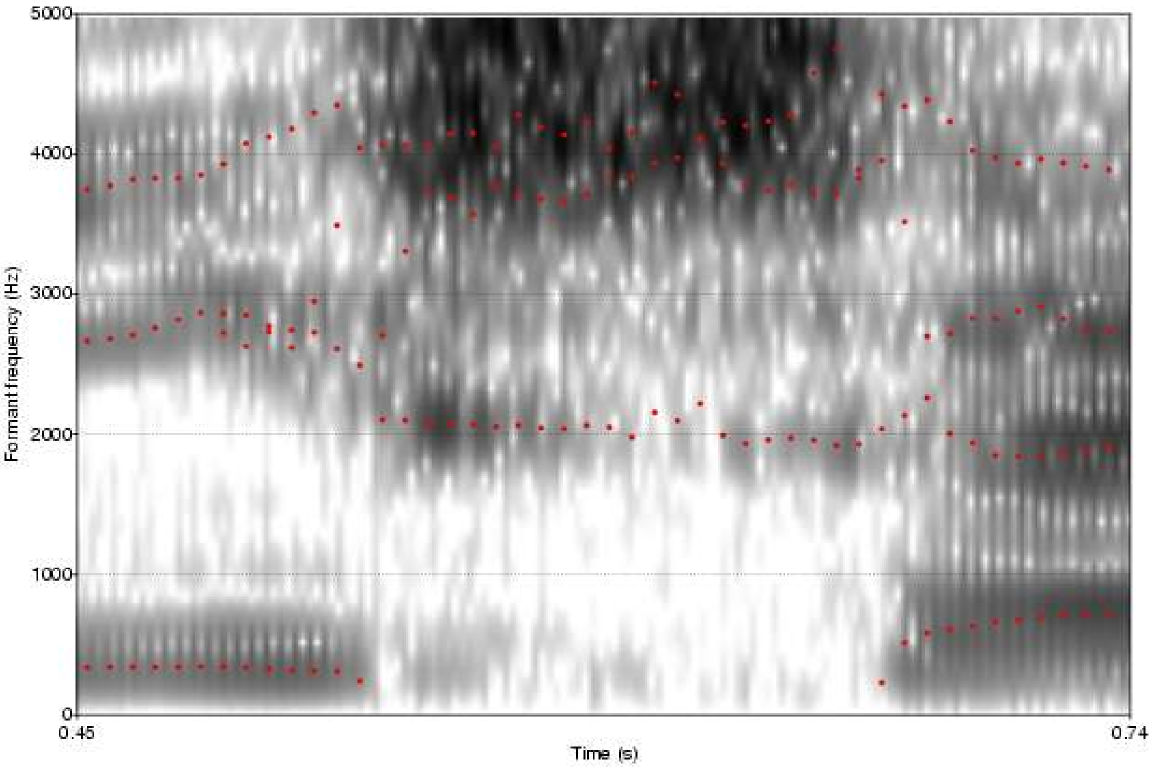


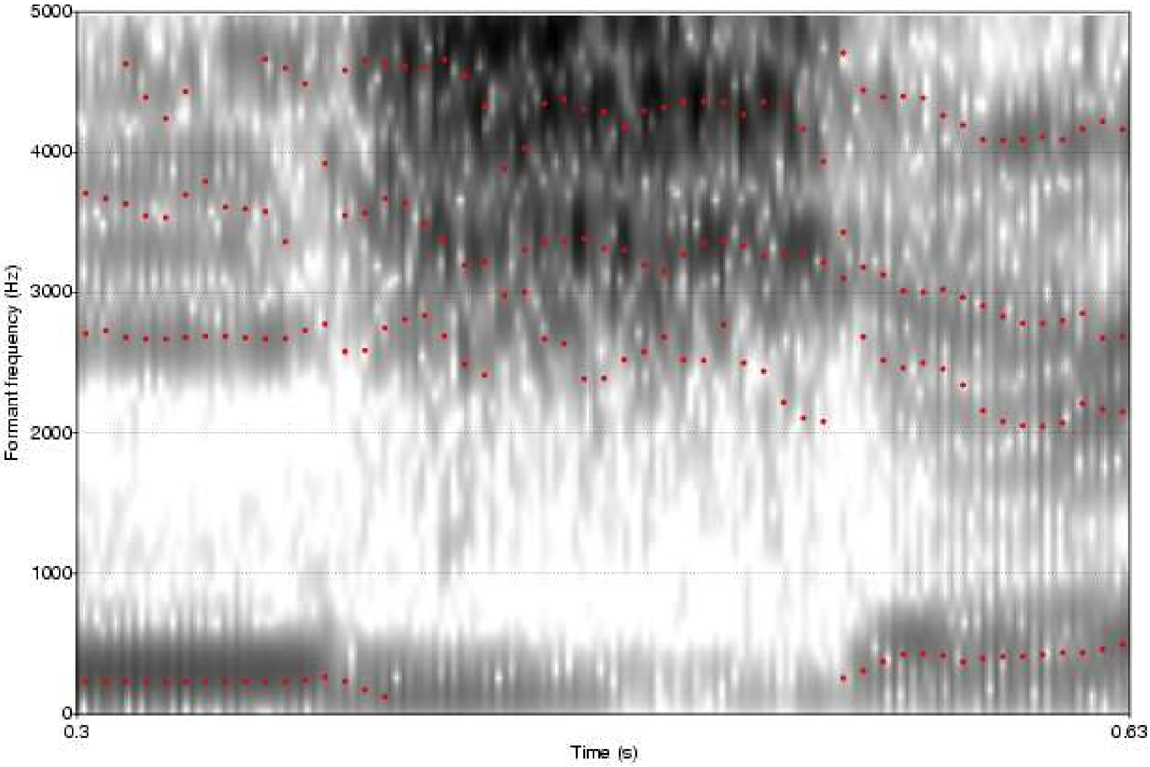
5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.33
0.74
Time (s)

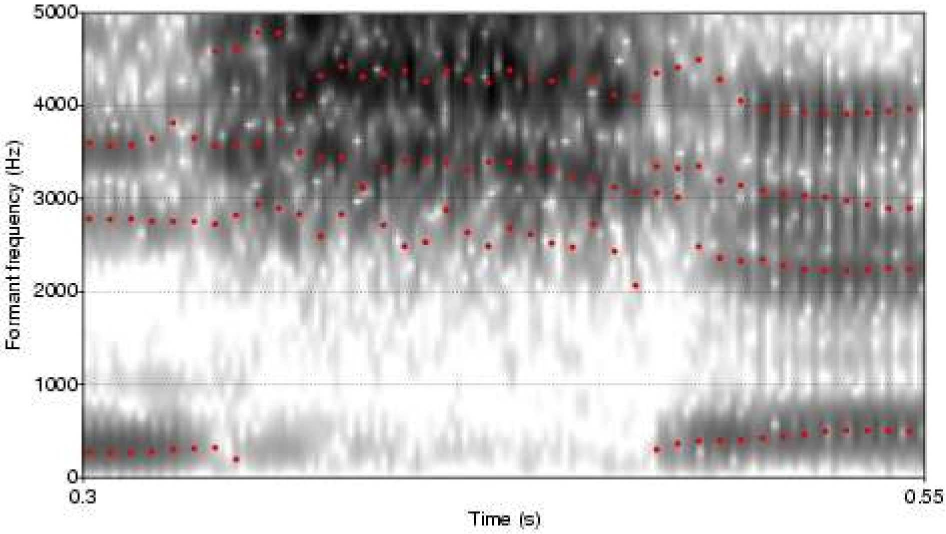


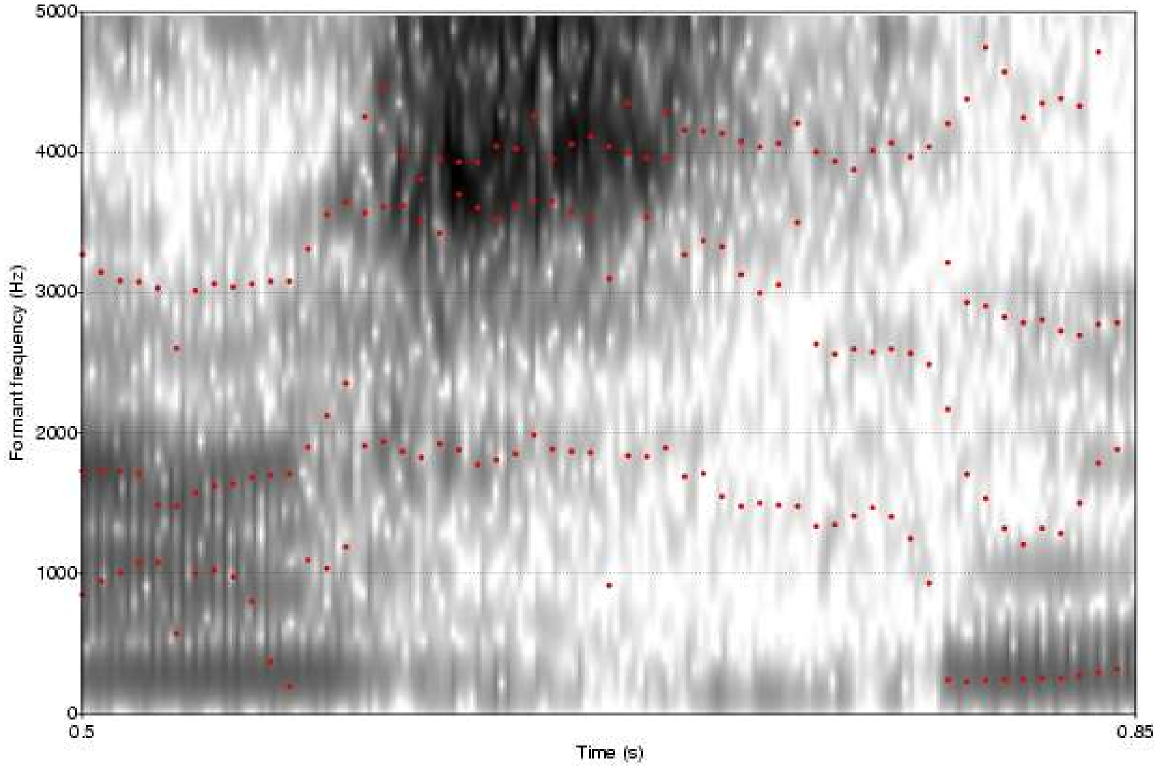
5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.29
0.6
Time (s)

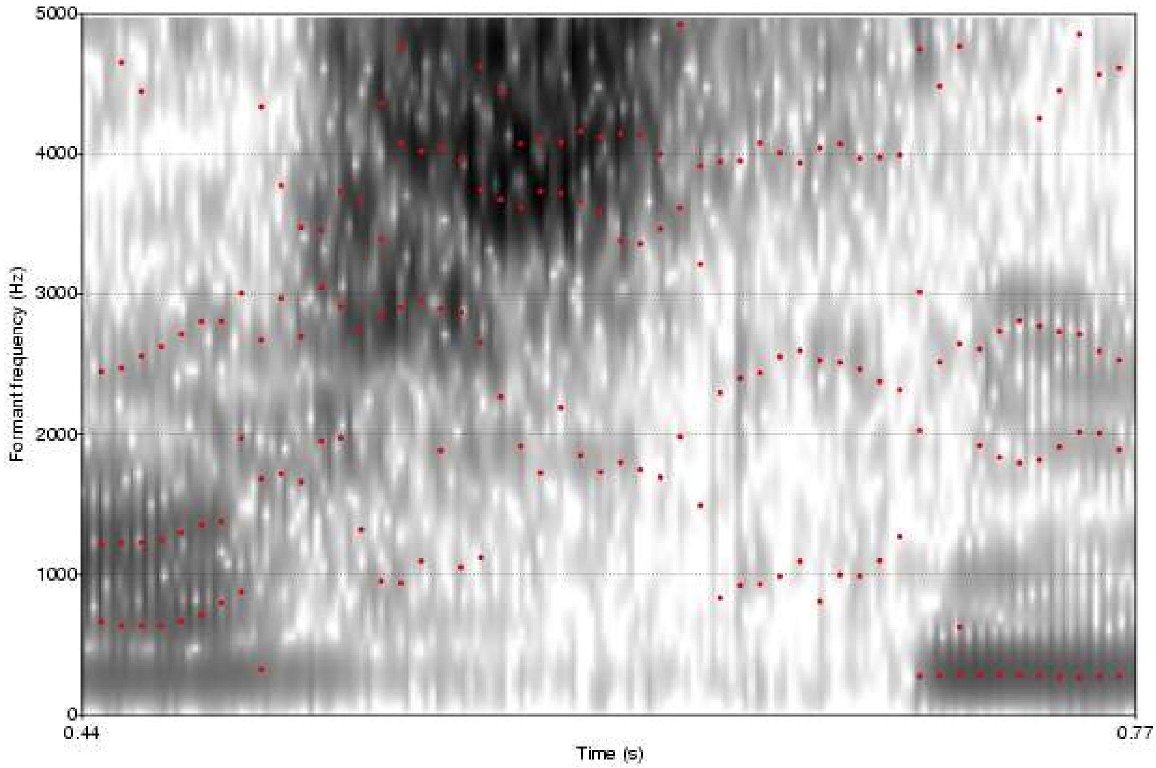

5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.5
0.81
Time (s)


5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.45
0.74
Time (s)


5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.3
0.63
Time (s)


5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.3
0.55
Time (s)


5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.5
0.85
Time (s)


5000
4000
3000
2000
1000
0
Formant frequency (Hz)
0.44
0.77
Time (s)