Acoustic analysis and sociolinguistic aspects of recent developments in Polish sibilants

Marzena Żygis
Zentrum für Allgemeine Sprachwissenschaft, Berlin
Humboldt-Universität zu Berlin

Bartłomiej Czaplicki
University of Warsaw

Daniel Pape
IEETA
University of Aveiro, Portugal

Introduction:
Articulation & Perception
of SIBILANTS

Phonetics:
Standard Polish /ɕ/ vs. new variant

Sociolinguistics:
Expressive Palatalization

Conclusions

Sibilants

- systematically resist coarticulation
  (Recasens 2007)
- are reluctant to undergo lengthening effects
  (Fougeron 2001, Fougeron & Keating 1997)
- highly prominent in perception
  (Ladefoged 1997)

→ sibilants do not like undergoing changes

EPG-based DAC model

DAC= Degree of Articulatory Constraint

(Recasens, Pallares & Fontdevila 1997)

- spatio-temporal patterns of tongue-dorsum and tongue-front coarticulation
- constriction location & width
- C-to-V & V-to-C effects

DAC values

	Tongue front		Tongue dorsum
	Constriction location	Constriction width	Position
n	Low		Low
dark l	Low		Low
s	High	High	High
trill r	High		High
ʃ	High	Low	Low
ʎ	Low		Low
k	Low		Low

s > trill r > ʃ > n, dark l, ʎ, k

Recasens (2004:436)

Effects of prosody (Fougeron 2001)

- Articulatory and acoustic properties at the initial boundaries of prosodic constituents

Intonational Phrase

↓

Accentual Phrase

↓

Word

↓

Syllable

linguopalatal contact
and duration of /t, n, k, l, **s**, i, ã/

Effects of prosody (Fougeron 2001)

Intonational Phrase

↓

Accentual Phrase

↓

Word

↓

Syllable

Increase from lower to higher position but less variation in the production of /**s**/

Linguo-palatal contact

Progressive lengthening from Syllable to Accentual Phrase but not always in case of /**s**/

Duration

→ prosodic position clearly affects the articulation of the stops and the lateral but less so /**s**/

Perceptual saliency

- Sibilants are better perceived than other fricatives; in contrast to [f v θ ð], they are not confused in perceptual experiments.

  (Harris 1958, Balise & Diehl 1994, Lisker 2001)

- "the well attested salient auditory characteristics [of sibilants] are the basis for the natural class "

  (Ladefoged 1997: 614)

→ sibilants do not like undergoing changes

Polish sibilant system

- alveolars: s z t͡s t͡z
- retroflexes (post-alveolars): ʂ ʐ t͡ʂ d͡ʐ
- alveolo-palatals: ɕ ʑ t͡ɕ d͡ʑ

- → an articulatory study suggests that only alveolo-palatals meet the criteria of palatalized consonants.

  (Rochoń and Pompino-Marschall 1999)

Current variation in Warszawa

Recordings of the current change:

New variant	[s sʲ ʂ]	[t͡sʲasnɔ]	[kasʲa]
Standard Polish	/s ɕ ʂ/	/t͡ɕasnɔ/ ‘tightly‘	/kaɕa/ ‘Kasia‘

Experimental evidence

0 Hypothesis:

Standard Polish /ɕ/ and /t͡ɕ/ do not differ from the new variants in the pronunciation of young female students in Warsaw

- Subjects

16 native speakers of Standard Polish from Warsaw, aged 19-23

- Phonemes

fricatives /s ɕ ʂ/

affricates /t͡s t͡ɕ t͡ʂ/

Material

- Words in controlled sentences:

- bisyllabic words stressed on the first syllable

- two contexts:

  word-initial _a

  word-medial a_a

- put into frame sentence :

  Powiedziala xxx do ciebie

  She said xxx to you

- 5 repetitions

- Words embedded in a text

- 13 sentences

- 3 repetitions

Acoustic parameters

- Using multitaper spectra (Thomson 2000, Lousada et al. 2012), we investigated the following acoustic parameters:

a. the highest spectral peak of the complete spectrum (0-16kHz)

b. the highest peak in the frequency range of 3-6kHz

c. the spectral moments according to the Praat formula (v. 5.2): Centre of gravity (COG), standard deviation of the spectrum (STD), skewness, and kurtosis

d. the spectral moments (M1, M2, M3, M4) according to Forrest et al. (1988)

e. the spectral slopes m1 and m2 (Jesus & Shadle 2002)

Acoustic parameters

f. Duration

g. The formants F1, F2, and F3 of the vowels preceding and following the consonant (endpoint frequencies)

h. The formant range of F1, F2, and F3 of the vowels preceding and following the consonant

- resampling to 11kHz,
- LPC spectra,
- window duration 0.0256 s,
- a prediction order of 13.

Statistics

LME models for the dependent variables (a-g)

Fixed effects

- VARIANT (Standard Polish, new variant),
- SPEECH STYLE (word, text),
- POSITION (word-initial, word-medial).

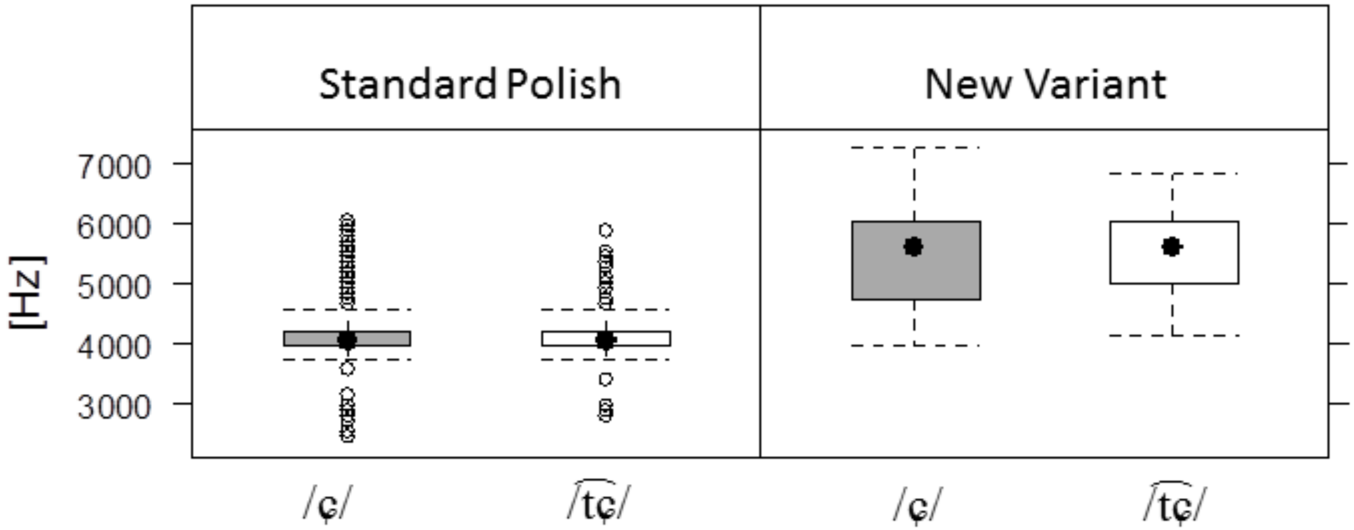
Random effects

- SUBJECT and ITEM,
- (random speaker-specific slopes for VARIANT)

All analyses in R environment

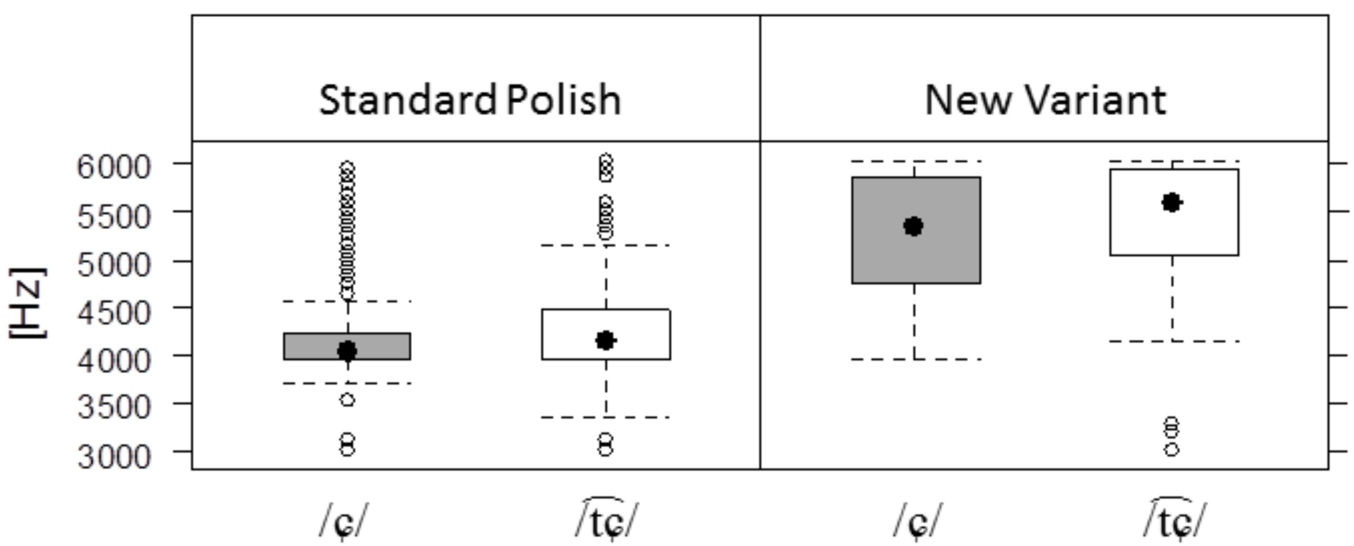
Peaks

20Hz-11kHz



/ɕ/: t=6.667
pMCMC = 0.0001
/t͡ɕ/: =6.289
pMCMC = 0.0001

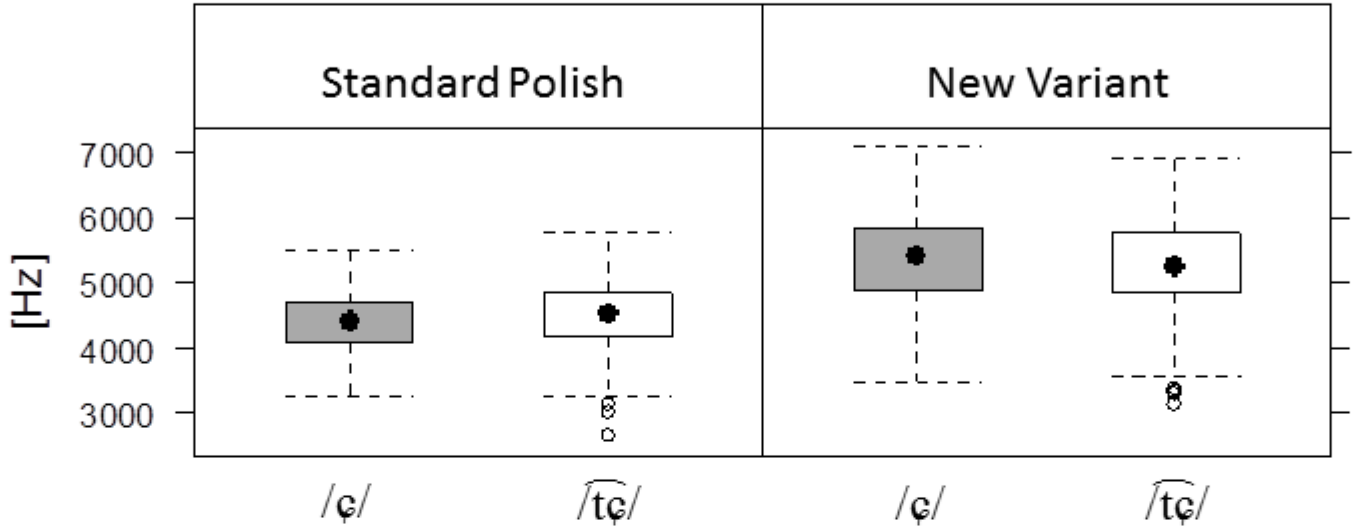
3-6kHz



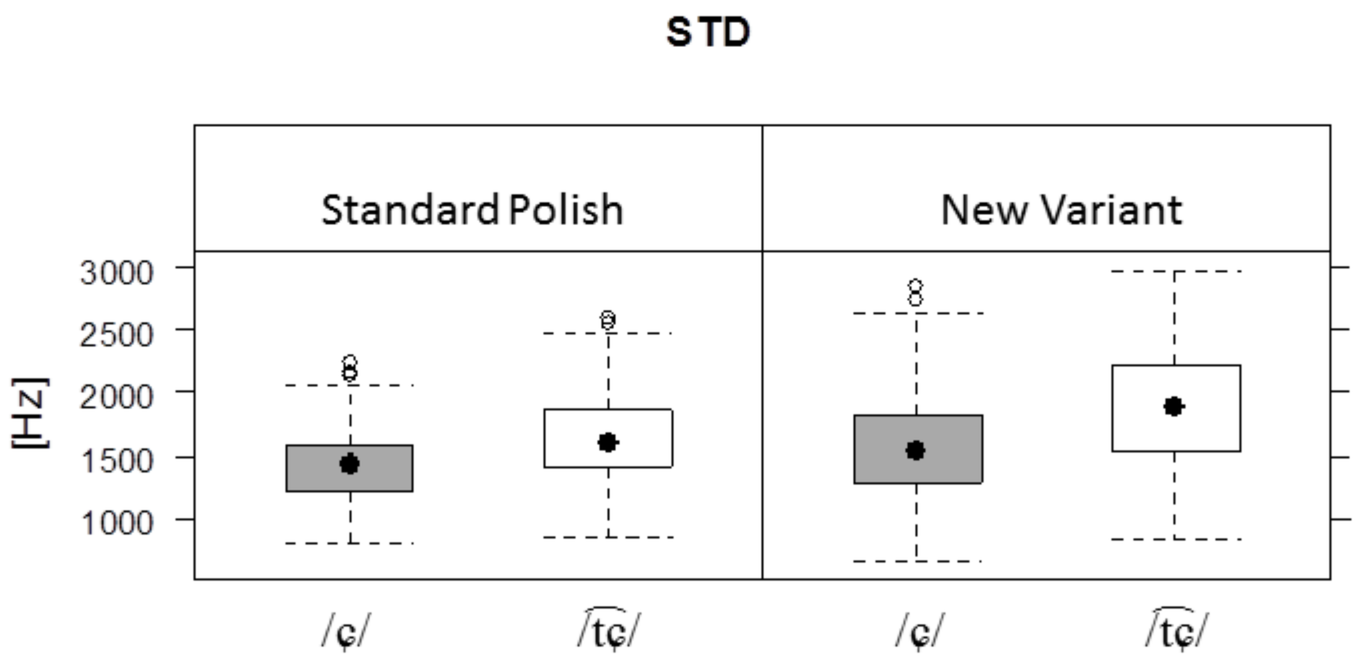
/ɕ/: t =7.208
pMCMC = 0.0001
/t͡ɕ/: t= 6.282
pMCMC = 0.0001

COG & SD

COG

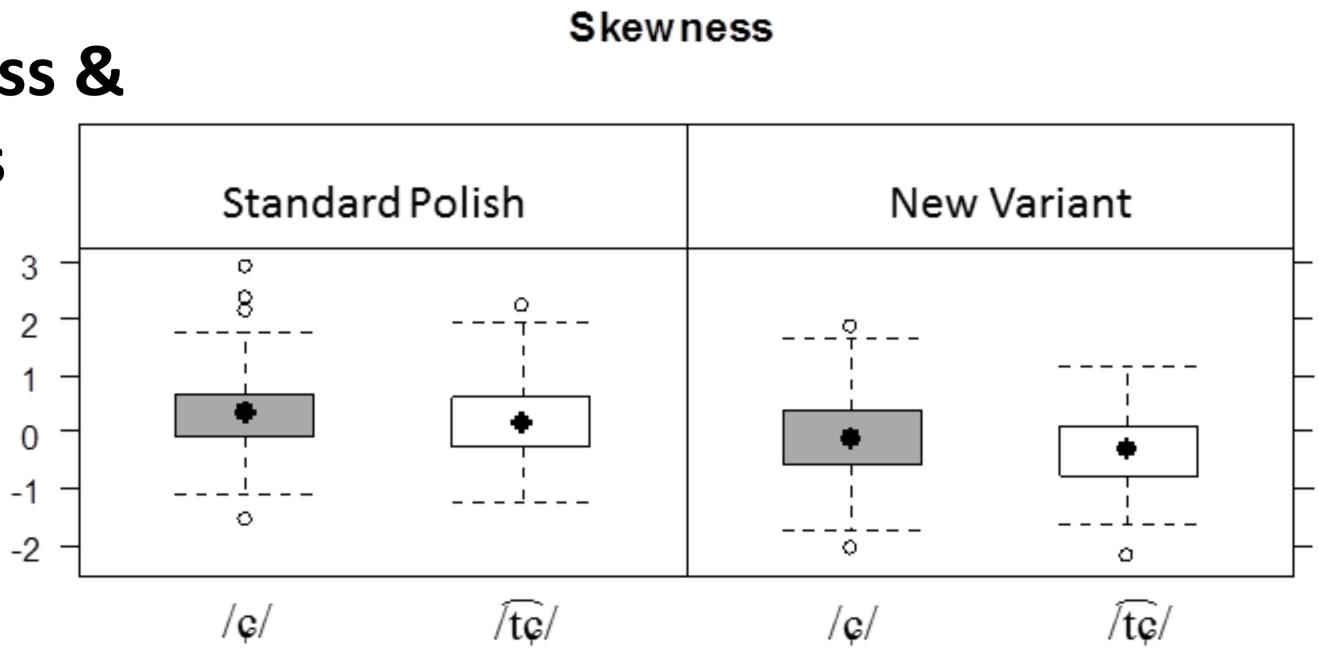


/ɕ/: t = 5.183
pMCMC = 0.0001
/t͡ɕ/: = 3.563
pMCMC = 0.0001

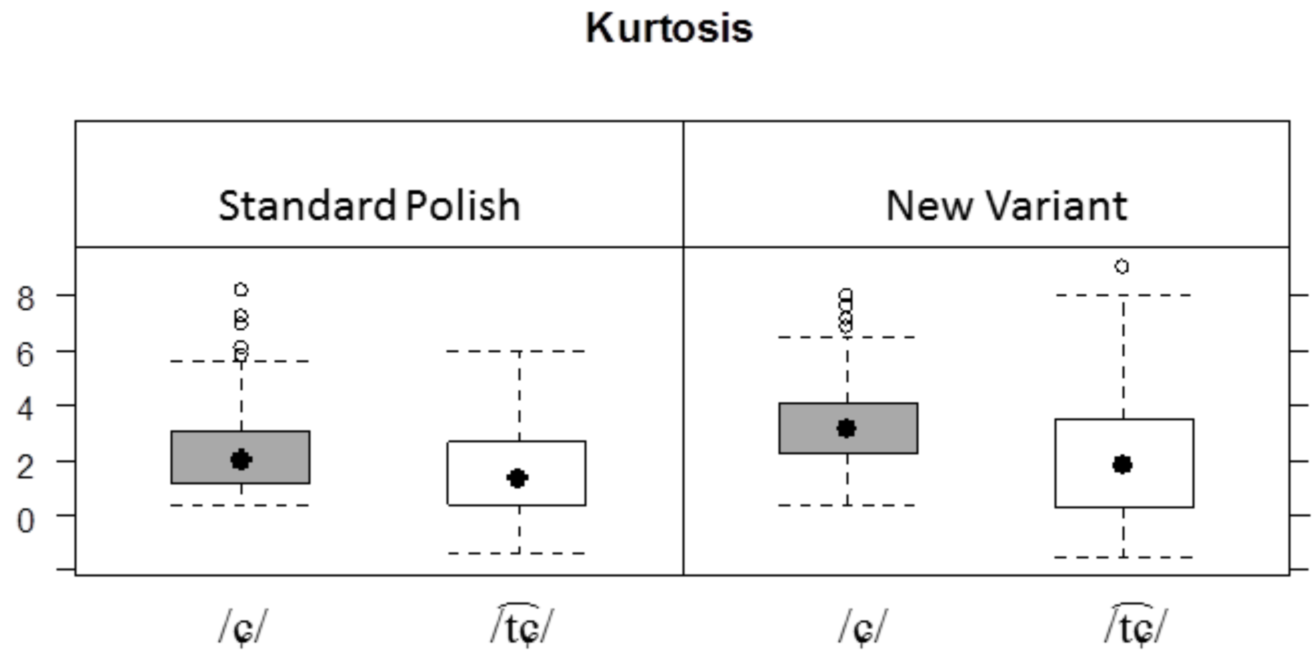


/t͡ɕ/: t = 1.725
pMCMC = 0.039

Skewness & Kurtosis

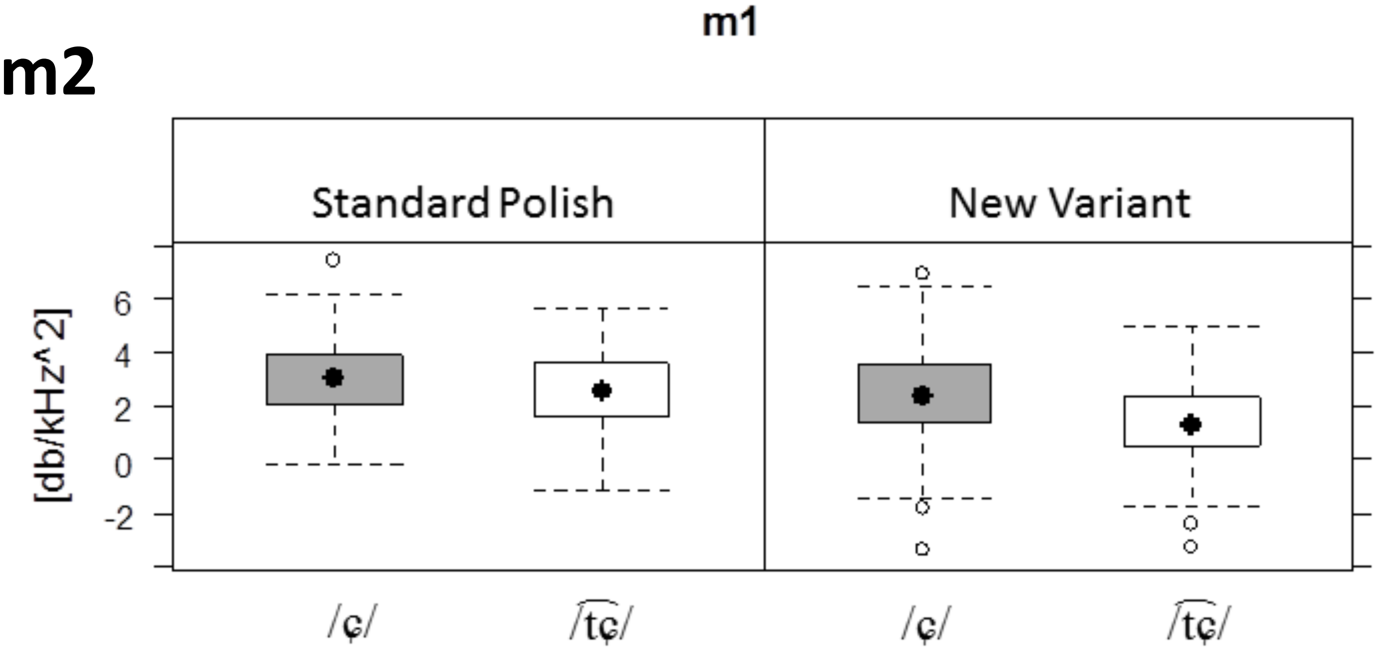


/ɕ/: t = -1.959
pMCMC = 0.0288
/t͡ɕ/: t = -2.718,
pMCMC = 0.005

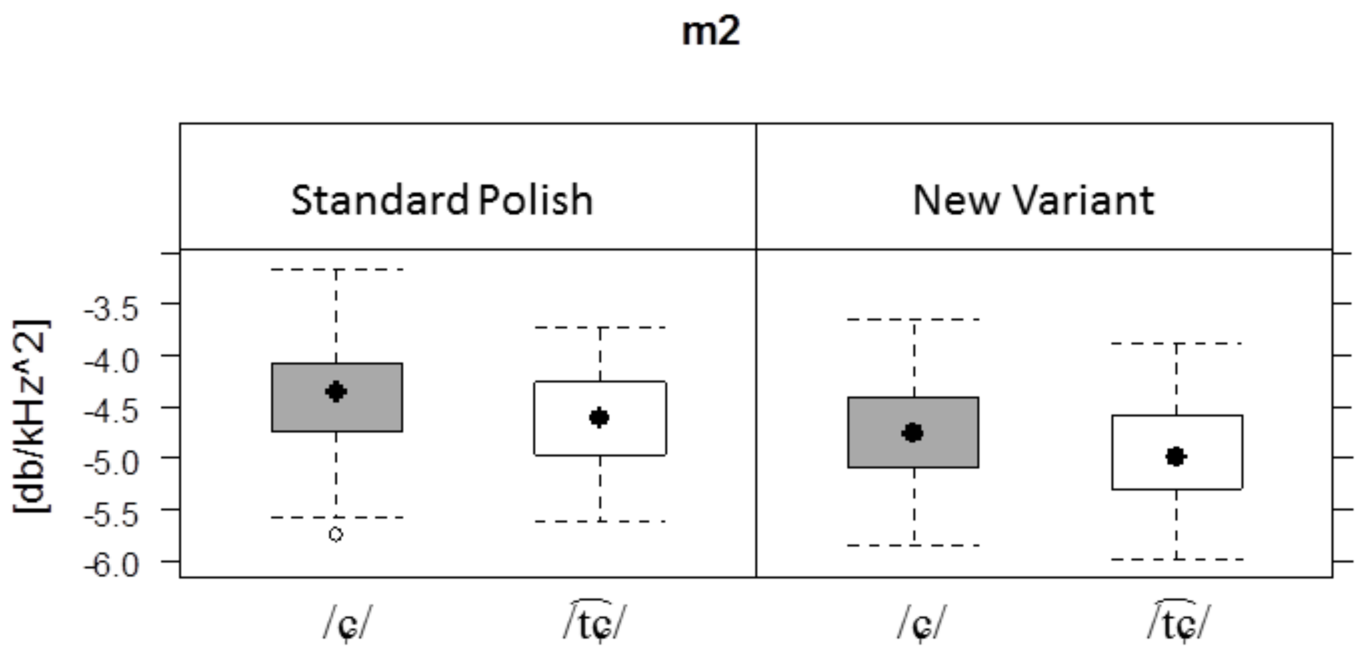


/t͡ɕ/: t=1.078,
pMCMC=.0104

m1 & m2

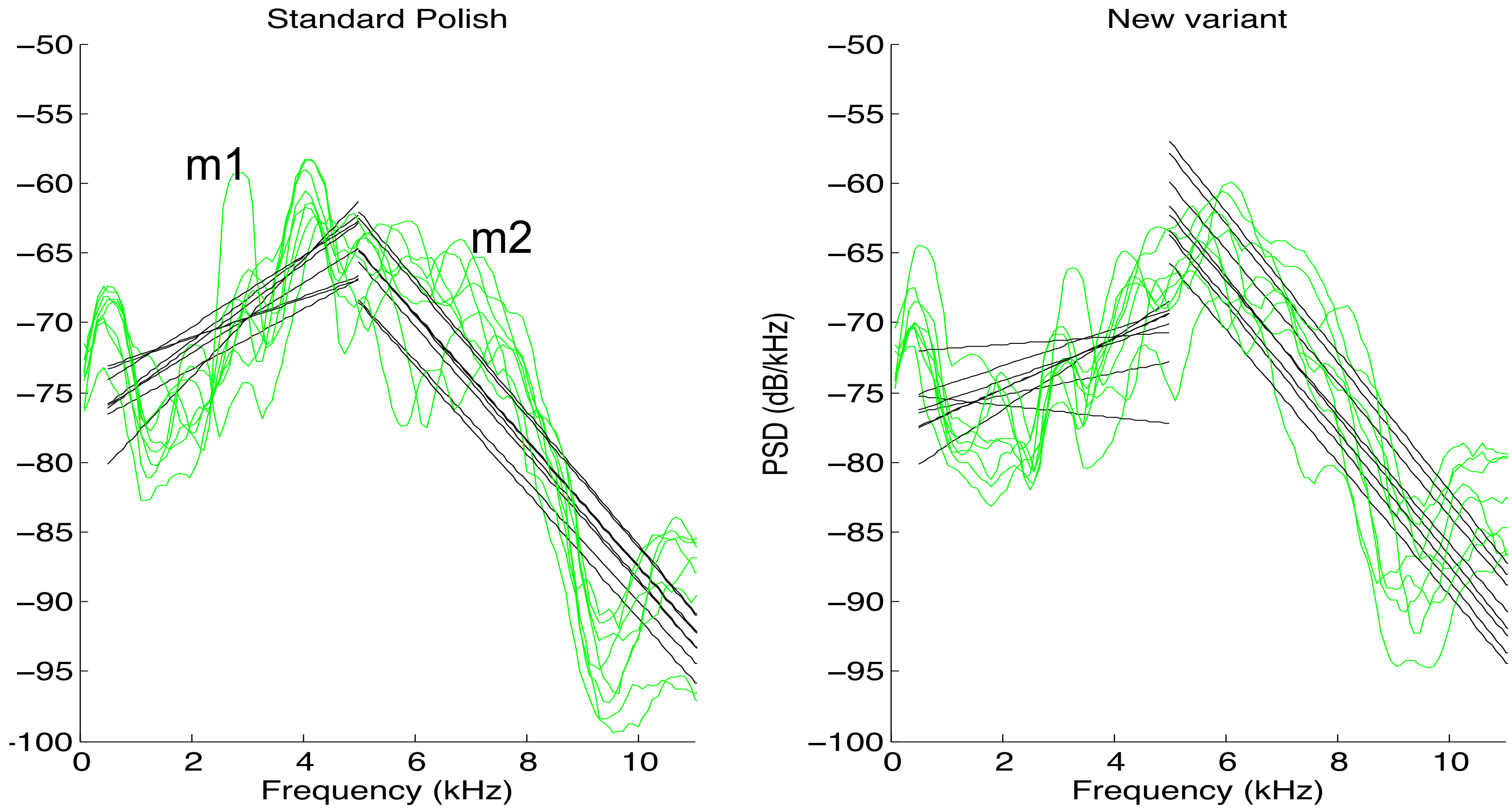


/t͡ɕ/: t= -2.363
pMCMC=0.003



/ɕ/: t = -3.72
pMCMC = 0.0008
/t͡ɕ/: t= -4.11
pMCMC=0.0001

m1 & m2 of /t͡ɕ/



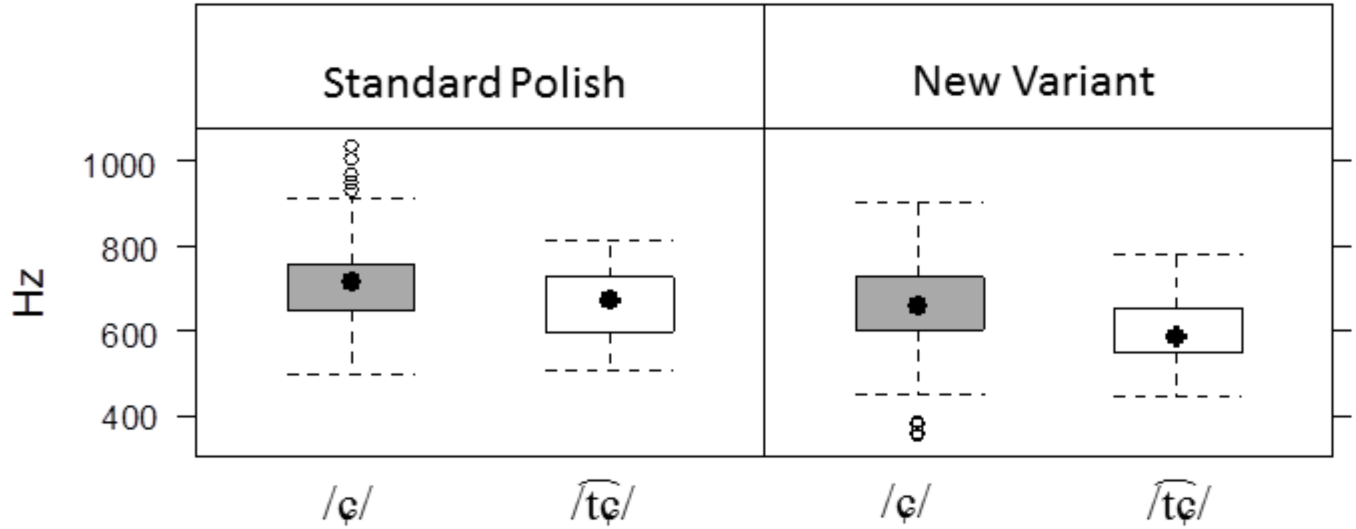
m1: /t͡ɕ/: t= -2.363
pMCMC=0.003

m2: /t͡ɕ/: t= -4.11
pMCMC=0.0001

→For 16 speakers

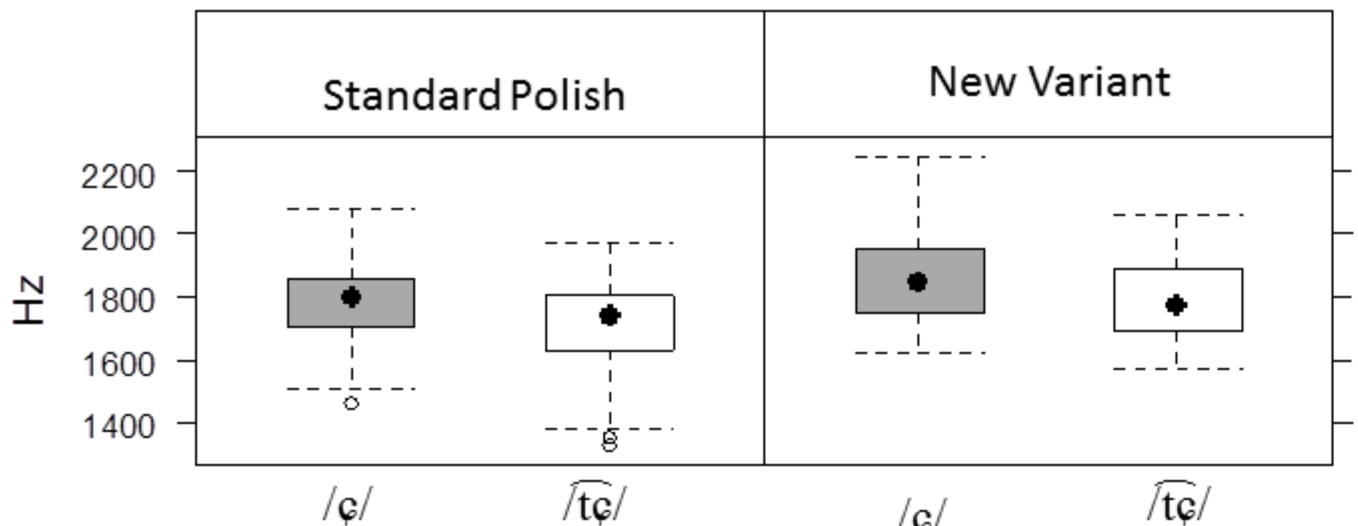
F1 & F2

F1 of the preceding vowel



/ɕ/: t = -1.581
pMCMC = 0.0494
/t͡ɕ/: t = -1.974
pMCMC = 0.031

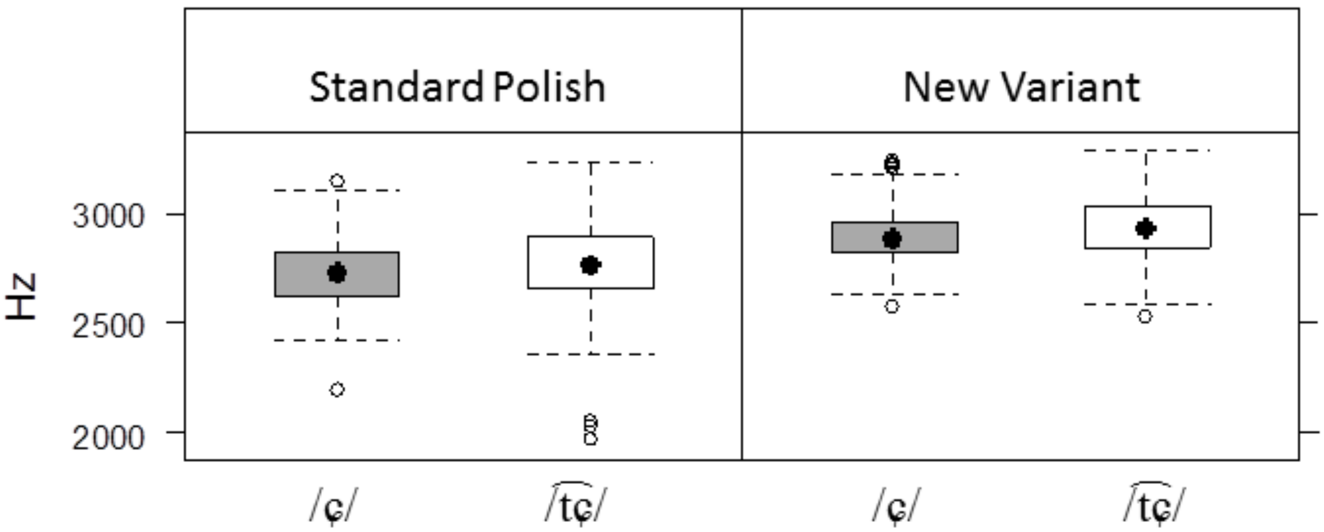
F2 of the preceding vowel



/ɕ/: t = 1.92
pMCMC = 0.0144
/t͡ɕ/::t = 2.17
pMCMC = 0.005

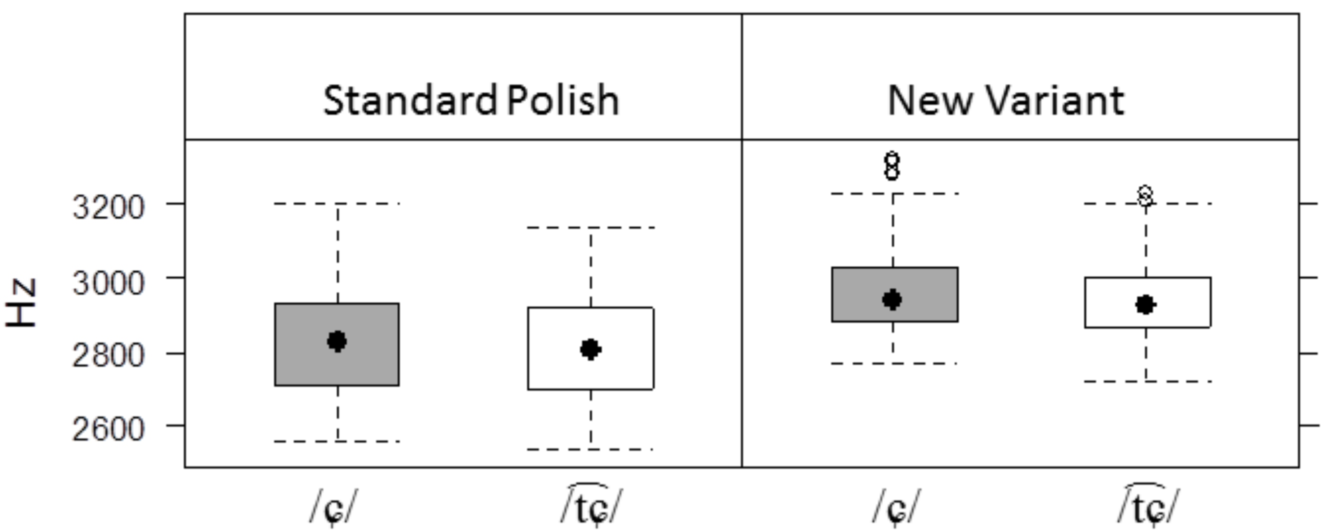
F3

F3 of the preceding vowel



/ɕ/: t = 3.78
pMCMC = 0.0001
/t͡ɕ/: t = 2.19
pMCMC = 0.0008

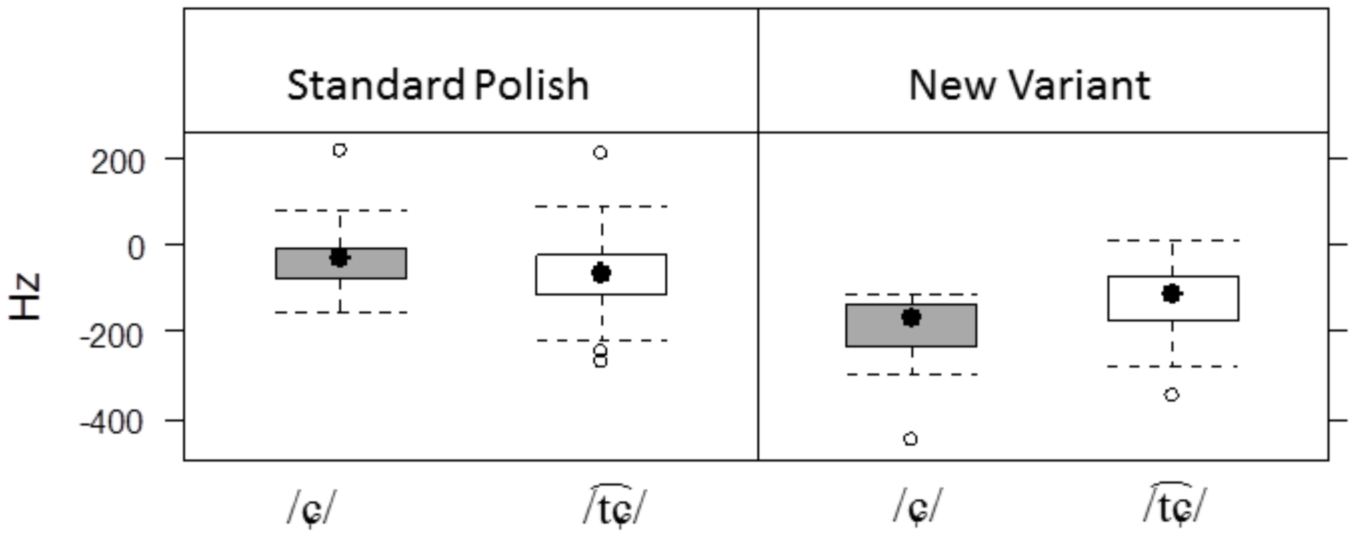
F3 of the following vowel



/ɕ/: t = 2.8
pMCMC = 0.0001
/t͡ɕ/::t = 2.33
pMCMC = 0.0002

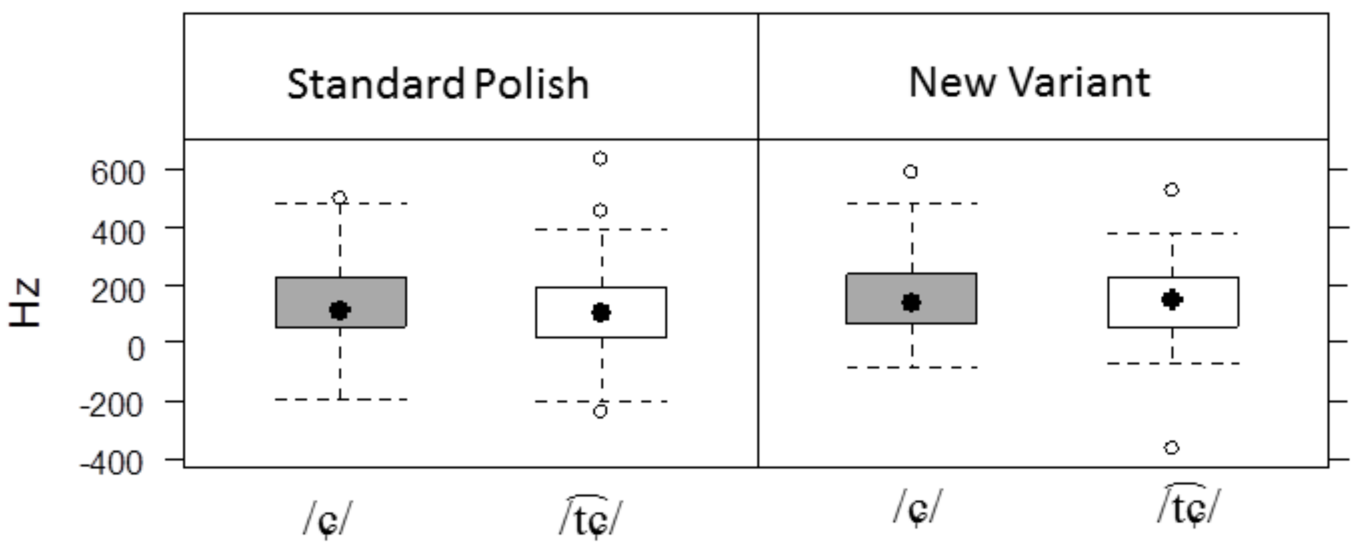
F1 & F2 rate

F1 rate of the preceding vowel



/ɕ/: $t = -12.934$
$pMCMC = 0.0001$
/t͡ɕ/: $t = -2.011$
$pMCMC = 0.0218$

F2 rate of the preceding vowel



n.s

Summary

The new variants substantially differ from Standard Polish /ɕ/ and /t͡ɕ/ in spectral properties:

- Spectral peaks (New variant > St. Polish)
- Centre of Gravity (New variant > St. Polish)
- Standard deviation (New variant > St. Polish)
- Skewness (New variant > St. Polish)
- Spectral slopes: m1 (New variant > St. Polish)
- Spectral slopes: m2 (New variant < St. Polish)

In addition,

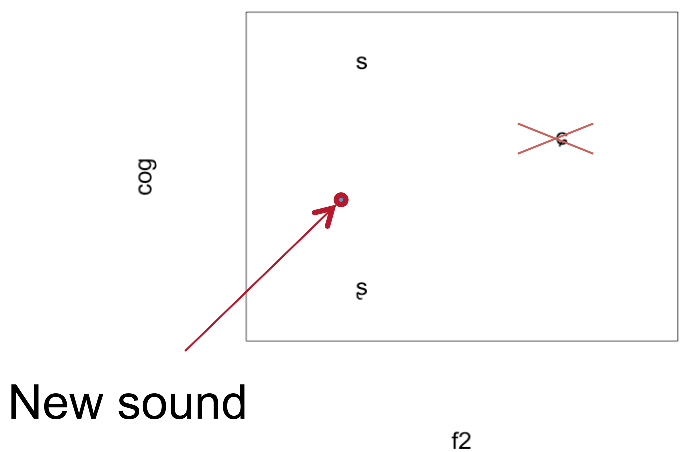
- F1 and F1 rate of the preceding vowel [a] (New variant < St. Polish)
- F2 & F3 of the preceding vowel [a] (New variant > St. Polish)
- F3 of the following vowel (New variant > St. Polish)

We are currently in the process of designing a perceptual experiment.

Some speculations about the triggers

1. Phonological phenomenon (perceptually optimal sibilants systems are strived for, Zygis & Padgett 2010)

- Less probable since the system would be not optimized (if we took only F2 and COG)



Some speculations about the triggers

2. COG values: the auditory distance between [ɕ] and [s] is greater than the distance between [ɕ] and [ʂ].

	ʂ	**ɕ**	**s**
COG (SP)	3500 Hz	4500 Hz	7500 Hz
COG (NV)	3500 Hz	5357 Hz	7500 Hz

The New Variant gives rise to more evenly dispersed contrasts in the system of sibilants (just looking at COG).

COG ʂ ɕ→ s

Questions

1. Why is the change initiated by women?
2. Why is the change initiated by young individuals?

The theory of dispersion of contrast does not deal with such issues.

→ sociolinguistics

Innovation vs. change

Innovation - "an act of an individual speaker, regardless of whether or not it later catches on in a speech community".

Change - "an innovation that has been widely adopted by members of such a community".

Janda and Joseph (2003)

Innovation: Expressive Palatalization

Concentration of energy in higher frequency regions that characterises certain palatal(ised) consonants and certain front vowels is universally associated with smallness.

- In animal communication, energy concentration in higher frequency regions (including high pitch) universally signals subordinacy, politeness and non-threatening attitude.
- Smaller vocal tracts emit higher pitch sounds.
- Males strive for low pitch. (Ohala 1994)

Expressive Palatalisation (EP) – an iconic relation between certain palatalised consonants and the meaning of "smallness", "childishness" and "affection".

- EP is a relatively common property of sound symbolism, diminutive morphology, hypocoristics and "babytalk". (Kochetov and Alderete 2011)

Japanese mimetics

[tʃoko-tʃoko] vs [toko-toko]	‘moving like a small child’ vs ‘trotting’
[katʃa-katʃa] vs [kata-kata]	‘the sound of keys hitting against each other’ vs ‘the sound of a hard object hitting the hard surface’
[p^{j}oko-p^{j}oko] vs [poko-poko]	‘hopping around in a childish bobbing motion’ vs ‘making holes here and there’

Japanese babytalk

[osarusan]	[oʧaruʧan]	‘monkey (honorific)’
[kuʦu]	[kuʧu]	‘shoe’
[tabemasuka]	[tabemaʧuka]	‘Will you eat?’
[omizu nominasai]	[omiʤu nominaʧai]	‘Drink your water!’

Polish hypocoristics

[magdalɛna]	[magduɕ̲a], [mad̲ʑ̲a]	‘Magdalena’, proper name
[marta]	[martuɕ̲a], [mart̲ɕ̲a]	‘Marta’, proper name
[mixaw]	[mixaɕ̲]	‘Michał’, proper name
[darjuʂ]	[daruɕ̲]	‘Dariusz’, proper name

Alveolo-palatals are commonly used for EP, the corresponding retroflexes [ʂ ʐ t͡ʂ d͡ʐ] do not appear in this function.

Providing motivation

The innovation involving alveolo-palatals, [ɕ tɕ] > [sʲ tsʲ], enhances the iconicity of these sounds by shifting energy weight into higher frequency regions.

This effect may be treated as a kind of exaggerated palatalisation or “overpalatalisation”.

Why has this innovation been introduced by adolescent girls?

➢The new variant has to do with transmitting the image of youth.

➢Such “overpalatalisation” effects often evoke an impression of the speaker’s childishness and immaturity.

Why is “overpalatalisation” not employed by adolescent boys?

➢The image that it conveys is not desirable to them.

Change – identity marker

Language change is most commonly observed among adolescents.

- Babies and young children are practically irrelevant for the study of change and its propagation.

This is rooted in social behaviour.

- The primary models for imitation for young children are their parents (adults).
- "At the preadolescent stage, we find the beginnings of a move from parent-oriented to peer-oriented networks". (Kerswill 1996: 196)

Change – identity marker

When children reach adolescence, their social networks are fully developed and their speech comes to resemble the speech of their peers.

At this age linguistic innovations initiated by some influential adolescents are the most likely to spread to their peers.

- Such linguistic traits become identity markers setting off a particular group from other, especially older speakers.

Change – identity marker

The newly introduced variants, [s^j ts^j], are employed exclusively by female adolescents.

- peer-oriented social pressures among members of this age group are the strongest.

Currently, we are witnessing an initial stage of a linguistic change.

- Whether the new variant catches on and is adopted by a wider community depends on the strength of social networks among adolescents and its perception as an identity marker within this group.

Parallel developments

A very similar phonomenon reported for Mandarin Chinese, Songyuan variety (Li 2005, Beckmann 2012).

New Variant	[s sʲ ʂ] = young women
	↑
Songyuan variety	/s ɕ ʂ/

Conclusions

1. The new variants substantially differ from Standard Polish /ɕ/ and /t͡ɕ/ (spectral peaks, centre of gravity, standard deviation, skewness, spectral slopes: m1, m2, as well as certain formant frequencies of the preceding and following vowels).
2. Energy concentration in higher frequency regions enhances the cues of iconic expressive palatalization and is used to connote the meaning of “youth” and “childishness”.
3. This innovation then spreads among adolescent girls and becomes an identity marker.
4. It remains to be seen whether this linguistic trait will establish its role as a marker of a Polish “feminine accent” in the future.

Senks!

Acknowledgments

- **We especially would like to thank Luis M.T. Jesus for the scripts on Multitaper and spectral slope analysis.**

- This research was supported by Federal Ministry of Education and Research (grant 01UG0711) to Marzena Zygis. It was also partially funded by FEDER through the Operational Program Competitiveness Factors - COMPETE and by Portuguese National Funds through FCT - Foundation for Science and Technology in the context of the project FCOMP-01-0124-FEDER-022682 (FCT reference PEst-C/EEI/UI0127/2011) to IEETA, as well as the post-doctoral fellowship from FCT (Portugal) grant SFRH/BPD/ 48002/2008 to Daniel Pape.

Multitaper spectra

- MT estimates can be used to generate spectrograms with much reduced ‘noise’ during fricatives (Blacklock & Shadle 2003).

- Spectral normalization (required prior to Spectral Moment calculation) is problematic, and generates high-tailed distributions. These result in both a high degree of correlation across moments, as well as a decrease in sensitivity in the higher-order moments.

- MT can accurately measure variations
  - Over short intervals during fricatives
  - Across tokens
  - Across speakers
  - Over long-term changes

- Multitaper offers promise for incorporating temporal variations in characterizations of fricatives

Table 1

Values for degree of articulatory constraint (DAC), front constriction location relative to the alveolar region, front constriction width, and tongue dorsum position for the Catalan consonants under analysis

	Tongue front				Tongue dorsum	
	Constriction location	DAC value	Constriction width	DAC value	Position	DAC value
n	Front	Low			Mid	Low
dark l	Front	Low			Low	Low
s	Back	High	Small	High	Low	High
trill r	Back	High			Low	High
ʃ	Back	High	Large	Low	High	Low
ʎ	Front	Low			High	Low
k	—	Low			High	Low