

# Coarticulation with alveopalatal sibilants in Mandarin and Polish: Phonetics or phonology?

Ivy Hauser  
*University of Minnesota*

## 1 Introduction

Previous work has shown that vowels following alveopalatal sibilants typically exhibit raised second formant (F2) values, typically attributed to coarticulatory vowel fronting (e.g. Stevens, 2004 in Mandarin; Bukmaier & Harrington, 2016 in Polish). This paper re-examines the palatalizing coarticulatory effects of the alveopalatal sibilant in Mandarin and Polish. While previous studies have focused on differences in F2 transitions or values at vowel onset, I find that the raised F2 values following alveopalatal sibilants frequently persist through the entire duration of following vowels in Mandarin. This raises the question of whether this is a phonetic coarticulation effect or a phonological assimilation effect. I review diagnostics for such a distinction and provide evidence from speech rate which suggests that the raised F2 effect should be analyzed as phonological assimilation in Mandarin, but phonetic coarticulation in Polish. These results have implications for phonological representations and perception in both languages.

**1.1 *Phonetics vs. phonology: Common diagnostics*** While the distinction between phonetics and phonology is tenuous, there are multiple diagnostics which have been used in the literature. The three main diagnostics we consider here are summarized below. This is not intended to be a complete list of diagnostics—there is a wealth of literature on the classification of “phonetic” vs. “phonological” effects (see e.g. Cohn (2007) for a review). I only intend to highlight these as commonly invoked diagnostics.

1. Gradience vs. categoricity (e.g. Chomsky & Halle, 1968)

Phonetic effects are expected to be gradient/continuous while phonological effects are expected to be categorical/discrete. Cohn (2007) argues that while this correlation is important, it is not as strong as is commonly assumed.

2. Extent of segmental effect (e.g. Keating, 1990)

Phonetic effects may only affect part of the segment while phonological effects are expected to affect the entire segment. This can be seen as a particular instance of the first diagnostic. Phonological effects should categorically affect the whole segment, creating discrete categories.

3. Variation with speech rate (e.g. Solé, 2007)

Phonetic (mechanical) effects should have fixed temporal extensions. Therefore, phonetic effects should not vary with speech rate. Phonological effects should have temporal extensions which do vary with speech rate.

**1.2 *Mandarin sibilants*** Many dialects of Mandarin contrast three voiceless sibilants. The non-alveopalatal sibilants have been characterized as having a variety of different places of articulation: dental, denti-alveolar, retroflex, laminal post-alveolar, and apical post-alveolar. Chang & Shih (2015) provides a review of these claims; it seems that some of this variation may be attributable to data collection in different

---

\* Many thanks to Diana Konarski, Zachary Sun, and Allison Chen, undergraduate research assistants who helped with collection and annotation of the data presented here. Thanks also to audiences at the University of Minnesota, the Annual Meeting on Phonology, Benjamin Munson, Kristine Yu, John Kingston, Joe Pater, and Gaja Jarosz for feedback and commentary on the project.

regions. In this study, I use the terms alveolar, retroflex and alveopalatal (followig e.g. Ladefoged & Wu, 1984; Duanmu, 2007; Chang & Shih, 2015). The consonant inventory of Mandarin is given in Table 1.

	Labial	Alveolar	Retroflex	Alveopalatal	Velar
Stop	p p <sup>h</sup>	t t <sup>h</sup>			k k <sup>h</sup>
Affricate		ts ts <sup>h</sup>	tʂ tʂ <sup>h</sup>	tɕ tɕ <sup>h</sup>	
Fricative	f	s	ʂ	ɕ	x
Nasal	m	n		ɲ	(ŋ)
Approximant		l			

**Table 1:** Consonant inventory of Mandarin (based on Duanmu, 2007)

Acoustically, the three sibilants have sometimes been described as having a three-way contrast in spectral center of gravity (COG; Lee, 1999; Lee-Kim, 2011; Kallay & Holliday, 2012). Other studies have reported a two-way COG contrast between the alveolar and the other two sibilants and an F2 onset contrast distinguishing the alveopalatal from the retroflex (Stevens et al., 2004). In perception, several studies have found the primary cue for the retroflex-alveolar contrast to be COG or the position of the lowest spectral prominence (Wu & Lin, 1989; Li, 2008; Chang, 2013). Li (2008) argues that COG is not sufficient to distinguish the alveopalatal from the other two sibilants and the primary cue distinguishing /ɕ/ is instead F2 onset of the following vowel.

There is an allophonic restriction on sibilants requiring [ɕ] before high front vowels (Duanmu, 2007; Lin, 2014).<sup>1</sup> Because of this positional neutralization, some have argued that the alveopalatals can be represented as underlying velars which become palatalized before high vowels (Wu, 1994), or that /s/ palatalizes to [ɕ] before [i] (Duanmu, 2007). However, all three sibilants contrast preceding the vowels [a] and [əu u] (Duanmu, 2007; Li, 2008; Lin, 2014). Therefore, all three sibilants are considered to be independent phonemes in many synchronic analyses (e.g. Li, 1999; Cheng, 2011) and are assumed to be distinct phonological categories in this study.

Some dialects of Manarin (Taiwan Mandarin and other southern varieties) merge the alveolar and retroflex sibilants to alveolar place. This merger is often partial: vowel context, contrastive focus, and social context all influence production of the contrast in speakers with the merger (Chang et al., 2013; Chang, 2013). None of our speakers appeared to exhibit the merger. Even if they did, neither the alveopalatal sibilant nor F2 of the following vowel are expected to be affected by the merger, so it is unlikely that the present results would be impacted.

**1.3 Polish sibilants** Many varieties of Polish contrast voiced and voiceless sibilants at three places of articulation: alveolar, alveopalatal, and retroflex (Dogil, 1990). The consonant inventory of Polish is shown in Table 2. Multiple studies (Nowak, 2006; Bukmaier et al., 2014) report a spectral center of gravity (COG) contrast between the dental fricative and the other two fricatives in production. Several studies have reported little difference in spectral center of gravity and other spectral measures of the frication noise (Żygis & Hamann, 2003; Bukmaier & Harrington, 2016; Lee-Kim, 2011). Instead, the alveopalatal and retroflex fricatives have been described as being distinguished by the transition of the second formant (F2) into the following vowel.<sup>2</sup>

When reporting on differences in vowel transitions, some quantify the vowel transition by the F2 difference at two vowel timepoints (Nowak, 2006; Chiu, 2009), and some quantify the transition by reporting the F2 value at vowel onset (Halle & Stevens, 1997; Kudela, 1968). Bukmaier & Harrington (2016) analyze onset of vowel transitions and show higher F2 values in the vowels following the alveopalatal, but these values showed some overlap with vowels following the retroflex. They report F2 trajectories between the onset and midpoint of the following vowel and argue that the raised F2 values following the alveopalatal are evidence for a coarticulatory palatalizing influence.

In perception, Nowak (2006) showed that frication noise alone is sufficient to categorize isolated

<sup>1</sup> It is worth noting that the Pinyin orthographic system does not accurately represent this distribution. Pinyin ‘xia’ and ‘xiu’ correspond to [ɕa əu] respectively, not [ɕia ɕiu] (Duanmu, 2007; replicated here).

<sup>2</sup> Żygis & Hamann (2003) also report data from a second speaker who has a three-way COG contrast between the sibilants.

	Labial	Alveolar	Alveopalatal	Retroflex	Palatal	Velar	Glottal
Stop	p b	t d				k g	
Fricative	f v	s z	ɕ ʐ	ʂ ʐ̥		x	h
Affricate		ts dz	tɕ dz	tʂ dz̥			
Nasal	m	n			ɲ		
Approximant		l ɹ			j	w	

**Table 2:** Consonant inventory of Polish (Padgett & Żygis, 2007)

fricatives for native speakers of Polish. However, it is possible the speakers are not interpreting the isolated fricatives as speech and therefore have enhanced discrimination. When the Polish fricatives were placed in a VCV context, removal of the formant transitions into the following vowel made the alveopalatal and retroflex fricatives confusable, indicating the primacy of F2 over COG as a cue to the contrast between /ç/ and /s ʂ/.

Sibilant mergers have also been observed in Polish, and it has been argued that the Polish three-way sibilant contrast is diachronically unstable. These arguments tend to center the retroflex as being particularly unstable and predict that it will merge with either the alveopalatal or the dental Bukmaier et al. (2014); Żygis et al. (2012). Mergers of both types have been reported in some nonstandard dialects of Polish (Żygis et al., 2012; Nowak, 2006; Bukmaier et al., 2014). None of the speakers analyzed here appeared to exhibit either merger.

## 2 Methods

### 2.1 Mandarin participants and stimuli <sup>3</sup>

11 Mandarin speakers were recorded. All speakers acquired Mandarin natively in China and relocated to the United States for college or high school. All speakers were between the ages of 18-30 and most speakers were undergraduate students. All recruitment materials (emails, sign-up info, etc.) were distributed in Mandarin orthography. All parts of the experiment were conducted in Mandarin by native speaker research assistants.

The stimuli in Mandarin were CV words and rare words which were intended to function as non-words. Because the Mandarin writing system is logosyllabic, developing new symbols for non-words presents several problems for participant reading. Instead of attempting to design new and orthographically natural characters, we used rare words with existing characters as “non-words”. Each stimulus was presented with the simplified Mandarin orthographic character and the pinyin script, a romanized quasi-phonemic orthographic system. With the pinyin presented alongside the logosyllabic characters, the participants were able to pronounce the intended stimulus even if they were unfamiliar with the word or Mandarin character. No participants self-reported trouble reading either orthographic system. The stimuli were read in the carrier phrase “wǒ bǎ X dú yī biàn” (‘I read X once’).

Mandarin stimuli were crossed according to the following factors: sibilant (3 levels: s ʂ ɕ) × vowel (3 levels: i a u) × word status (3 levels: high frequency/low frequency/non-word) × number of syllables (2 levels) × tone (4 levels). Not all factors could be fully crossed: there is a phonotactic restriction that requires the alveopalatal sibilant before [i], so the three sibilants are only fully crossed in the [a] and [u] contexts. Due to limitations of the lexicon, some of the tones are not fully crossed with all other factors. There were a total of 137 distinct sibilant stimuli. Additional stimuli with word-initial affricates and stops were included as fillers. Word-initial non-sibilant fricatives were not included in the task. Stop and affricate tokens were elicited as fillers along with the voiceless sibilants, neither of which are analyzed here.

**2.2 Polish participants and stimuli** The Mandarin data are compared with Polish data which come from a separate study of 3 native Polish speakers, all undergraduate students at the University of Massachusetts Amherst who had acquired Polish natively in Poland.

The stimuli were words and non-words where the onset consonant was a sibilant and the following vowel

<sup>3</sup> All materials (stimuli lists, data, code) are available in the public data archive for this project at <https://osf.io/2famr/>.

was one of [ɛ a ɔ]. Voiced and voiceless sibilants were elicited, but only the voiceless sibilants are analyzed here (Mandarin does not have a corresponding voiced series for comparison). Due to the rich inflectional system in Polish, the lexicon does not include many monosyllabic words. Therefore, the Polish stimuli included mono-, di-, and trisyllabic words.

As in the Mandarin experiment, the stimuli were cross-classified according to word frequency. In The words were classified into two frequency categories, high and low. We used the Polimorf lexicon (Wolinski et al., 2012) for selecting stimuli and verified native speaker frequency intuitions with orthographic frequency data.<sup>4</sup> All the stimuli were recorded in the carrier phrase: “Powiedziała X od razu” (‘She said X right away’). Stimuli were crossed according to: sibilant (6 levels) × vowel context (3 levels) × word status (3 levels: high frequency/low frequency/non-word) × number of syllables (3 levels). Due to gaps in the Polish lexicon, not all factors could be fully crossed<sup>5</sup> and the full stimuli set included 126 stimuli.

**2.3 Recording and data processing** The participants were all recorded in a sound-attenuated booth using Audacity software (Audacity Team, 1999-2014) with an M-Audio Fast Track Pro Mobile Audio Interface and a Shure SM10A head-worn microphone. The recordings were sampled at a rate of 44.1 kHz with a bit depth of 16.

The participants were presented with stimuli in the relevant orthography on a laptop computer inside the booth. They were asked to produce the phrases as naturally as possible. The research assistants were trained to give feedback which encouraged natural production.<sup>6</sup> The stimuli were recorded in four separate blocks, each with a different random order, totaling four repetitions of each stimulus for analysis.

After recording the participants completed a word frequency judgment task with the stimuli. This was to ensure that the frequency data matched the intuitions of the participants. The judgment task consisted of filling out a survey indicating degree of familiarity with each word and took between 2-5 minutes to complete.

**2.4 Data processing and analysis** The recordings were first scanned by the author and research assistants for speech errors. The recordings were forced aligned using the Montreal Forced Aligner (McAuliffe et al., 2017), which creates Praat (Boersma et al., 2001) textgrids marking boundaries at the word and segment level. The Mandarin data were aligned using a pretrained Mandarin model.<sup>7</sup> A new model was trained to align the Polish data.

A Praat script based on DiCanio (2013) was used to extract spectral moments of the fricatives and formant values of the following vowels. The spectral moments are not analyzed in this paper. The formants were estimated using the Burg method and extracted at 10 ms intervals throughout the duration of the vowel. Formant excursions greater than 1000 Hz over 10 ms were assumed to be tracking errors and were excluded.

### 3 Results

**3.1 Mandarin** As expected based on previous literature, we found consistent differences in onset F2 following the alveopalatal sibilant relative to the other sibilants. Somewhat less expected, we found that these differences frequently persisted throughout the entire duration of the vowel. For /u/, all speakers consistently used higher F2 following the alveopalatal sibilant throughout the vowel duration. For /a/, there was more between-speaker variation in extent of F2 raising—some speakers extended raised F2 throughout the entire duration of the vowel, while others extended raised F2 through 50-75% of the vowel. There were also vowel differences in within-speaker variation, with speakers generally exhibiting more within-category variation in F2 for sibilants preceding /u/. Data from representative speakers are shown in Figure 1. All figures were created using the ggplot2 package (Wickham, 2009) in R (R Core Team, 2013). These graphs show average F2 trajectories with Loess smoothing over normalized time.

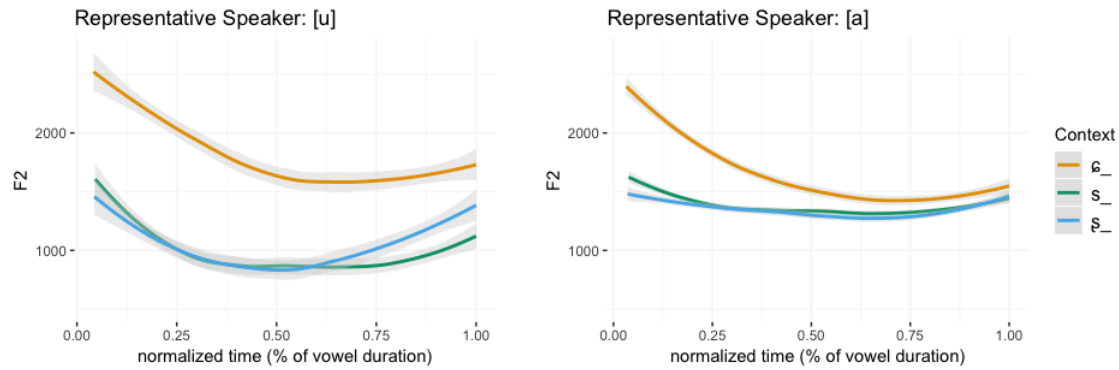
A mixed-effects linear regression was performed predicting F2 at vowel *offset* (as we are interested in whether or not the raising effect persists through the entire vowel). Fixed effects were preceding sibilant (C), vowel (V), and vowel duration (as an indicator of speech rate). Interactions between C×V and C×

<sup>4</sup> Thanks to Gaja Jarosz for providing this data and offering helpful commentary on the experimental design.

<sup>5</sup> Non-words were used, however, word status is still not fully crossed with all the other factors if there are lexical gaps.

<sup>6</sup> This included things like suggesting the participant speak as if they were talking to a friend and not giving a presentation, suggesting they say the phrase “in one breath” to discourage pausing before the stimulus, etc.

<sup>7</sup> Available at [https://montreal-forced-aligner.readthedocs.io/en/latest/pretrained\\_models.html](https://montreal-forced-aligner.readthedocs.io/en/latest/pretrained_models.html).



**Figure 1:** F2 trajectories in Mandarin: average time-normalized formant values for two representative speakers in two vowel contexts.

vowel duration were also included, along with random intercepts for speaker and word. The fixed effects are summarized in Table 3.

Effect	Estimate	Std. Error	df	t value	p
(Intercept)	1534.62	105.80	33.37	14.505	***
C-s	-616.55	179.46	25.55	-3.435	**
C-ʃ	-435.79	167.67	28.54	-2.599	*
V-a	-21.40	178.22	24.89	-0.120	
vowel.dur	1085.71	274.86	1521.98	3.950	***
C-s × V-a	333.60	254.53	17.97	1.311	
C-ʃ × V-a	258.98	236.82	18.41	1.094	
C-s × vowel.dur	1738.41	403.67	1547.68	4.307	***
C-ʃ × vowel.dur	996.73	410.38	1551.66	2.429	*

**Table 3:** Fixed effects table for mixed-effects linear regression with Mandarin data. Dependent variable is F2 value at vowel offset. Model intercept is /ɕu/.

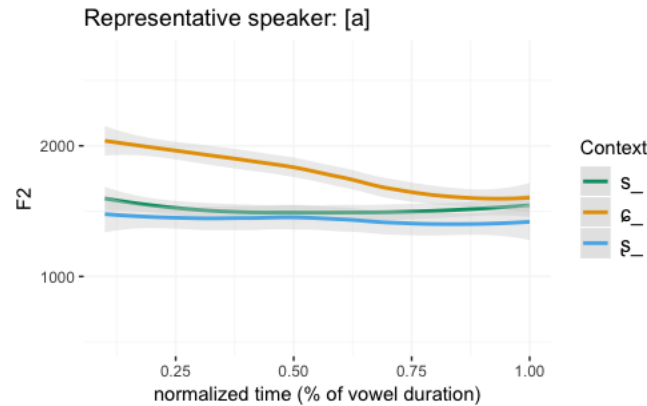
There is a significant effect of preceding C on F2 offset, with F2 higher after the alveopalatal [ɕ] relative to both [s ʃ]. There is also no significant effect of the vowel [a], and no significant interactions between [a] and the other sibilant contexts. Together, these results indicate that F2 values at vowel offset are significantly higher following alveopalatal [ɕ] in both vowel contexts.

There is also a significant effect of vowel duration for the intercept [ɕu] with a positive estimate. This indicates that F2 at vowel offset *increases* with vowel duration. The significant interactions between [s ʃ] and vowel duration indicate that this relationship also holds in the other sibilant contexts. Together, these results indicate that F2 values at vowel offset are significantly higher in longer vowels following all sibilants.

**3.2 Polish** As expected based on previous literature, the Polish speakers also produced differences in F2 transitions. While the speakers sometimes produced F2 differences that persisted through the entire duration of the vowel, the results are much less consistent (across speakers and tokens) relative to the Mandarin data.

Data for [a] from a representative Polish speaker are shown in Figure 2. For this speaker (as with the other Polish speakers), the raised F2 of [a] following [ɕ] typically does not persist through the entire vowel. This is somewhat similar to the Mandarin [a] data, where some speakers also did not extend the raised F2 throughout the vowel duration. The Polish data differs from the Mandarin data in that there were no vowel contexts where speakers consistently raised F2 for the entire vowel duration.

A mixed-effects linear regression was performed on the Polish data with the same effect structure used for the Mandarin data (described in §3.1). There were no significant effects of vowel quality, preceding



**Figure 2:** F2 trajectories in Polish: average time-normalized formant values for [a] from a representative speaker.

sibilant, or vowel duration on F2 offset.

#### 4 Discussion

In Mandarin, our results show that the influence of the alveopalatal sibilant on the following vowel is not merely in the transitions to the vowel; vowels following the sibilant have a higher F2 throughout the duration of the vowel, even at vowel offset. I argue that these data suggest a phonological analysis of vowel fronting following alveopalatal sibilants in Mandarin.

Solé (2007) proposes that phonetic or “mechanical” effects should have temporal extensions which are fixed and independent of speech rate. Therefore, if raised F2 were a purely mechanical effect, we would see raised F2 for a fixed period of time across all vowel durations. In that case, we would expect a significant negative effect of vowel duration on F2 offset. However, we found a significant positive effect of vowel duration across all sibilant contexts, including [ʃ]. F2 values following [ʃ] are higher for longer vowels at vowel offset. This is the opposite of what would be predicted if F2 raising were a purely mechanical effect.

The fact that F2 offset is higher in longer vowels for all sibilant contexts is likely due to increased anticipatory coarticulation with the following segment [d]. Despite this anticipatory raising in the other sibilant contexts, F2 at vowel offset is still significantly higher following [ʃ] in both [a] and [u]. In longer vowels, speakers raise F2 offset in all sibilant contexts while maintaining the vocalic distinction between [ʃ\_] and [s ʃ\_]. If the vowel fronting following [ʃ] were purely mechanical/phonetic, it is not likely that we would see this distinction maintained during increased anticipatory coarticulation at vowel offset.

A phonological fronting process in Mandarin could be analyzed with the feature [+front] spreading rightward from the alveopalatal sibilant to following vowels. However, some phonological analyses of Mandarin assume that [ʃ] is always derived from underlying [si] sequences. For example, [ca] sequences are derived from underlying /siu/ via regressive assimilation of [+front] (Duanmu, 2007). It is difficult to maintain this analysis in light of the present results, which suggest the vowel following the alveopalatal sibilant is also [+front]. This would require the underlying /i/ to trigger bidirectional assimilation to both the preceding /s/ and the following /a/. If we instead assume an underlying [ca] sequence, we can easily analyze the fronting as assimilation where [+front] spreads from the sibilant to the following vowel.<sup>8</sup>

In Polish, our results show that the coarticulatory vowel fronting following alveopalatal sibilants does not have a significant effect on F2 at vowel offset—the effect does not extend throughout the vowel duration. This,

<sup>8</sup> I assumed this analysis when placing [ʃ] in the consonant inventory chart for Mandarin in Table 1. This departs from Duanmu (2007), who follows the analysis that [ʃ] is derived from underlying [si] sequences and does not include [ʃ] in the inventory chart.

combined with the lack of any speech rate effects, suggest a phonetic analysis of alveopalatal coarticulation in Polish. In our data, Polish speakers typically only exhibit raised F2 for a portion of the vowels following [ɕ].

Previous work on perception has found that Polish speakers were only able to correctly identify sibilants when the appropriate following vowel transitions were present Nowak (2006). The present results suggest the potential for a stronger perceptual effect in Mandarin. Because F2 differences following alveopalatal sibilants typically extend throughout the entire vowel duration, Mandarin speakers may be able to identify the preceding fricative from only hearing the following vowel, or even only the offset of the following vowel. Further perceptual work will need to be done to determine exactly how listeners use the vocalic information for sibilant identification.

## 5 Conclusion

In Mandarin and Polish, vowels following alveopalatal sibilants exhibit raised F2, though the patterns differ between the two languages. I argue that F2 raising is due to phonetic coarticulation in Polish—raising only affects the vowel transitions and not the whole segment, and extent of raising does not appear to vary with vowel duration. However, in Mandarin, I argue that F2 raising is due to phonological assimilation—raising affects the whole segment, regardless of vowel duration. Further work will need to be done to determine the effects of raised F2 on perception, specifically raise whether listeners might use the entirety of vocalic information for sibilant identification.

## References

- Audacity Team (1999-2014). *Audacity(R): Free Audio Editor and Recorder*. <http://audacity.sourceforge.net/>.
- Boersma, Paul et al. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5:9/10, 341–345.
- Bukmaier, Véronique & Jonathan Harrington (2016). The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *Journal of the International Phonetic Association* 46:3, 311–329.
- Bukmaier, Véronique, Jonathan Harrington, Ulrich Reubold & Felicitas Kleber (2014). Synchronic variation in the articulation and the acoustics of the Polish three-way place distinction in sibilants and its implications for diachronic change. *Fifteenth Annual Conference of the International Speech Communication Association*.
- Chang, Yung-Hsiang (2013). *Variability in cross-dialectal production and perception of contrasting phonemes: the case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin*. Ph.D. thesis, University of Illinois at Urbana-Champaign.
- Chang, Yung-Hsiang Shawn & Chilin Shih (2015). Place contrast enhancement: The case of the alveolar and retroflex sibilant production in two dialects of Mandarin. *Journal of Phonetics* 50, 52–66.
- Chang, Yung-Hsiang Shawn, Chilin Shih & Jont B Allen (2013). Dialectal variation in the perception of phonological contrasts. *Proceedings of the International Conference on Phonetics of the Languages in China*, 115–118.
- Cheng, Chin-Chuan (2011). *A synchronic phonology of Mandarin Chinese*, vol. 4. Walter de Gruyter.
- Chiu, Chenhao (2009). Acoustic and auditory comparisons of Polish and Taiwanese Mandarin sibilants. *Canadian Acoustics* 37:3, 142–143.
- Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. Harper & Row.
- Cohn, Abigail C (2007). Phonetics in phonology and phonology in phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16, 1–31.
- DiCano, Christian (2013). Time averaging for fricatives 2.0. Praat script published online.
- Dogil, Grzegorz (1990). Hissing and hushing fricatives: A comment on non-anterior spirants in Polish. *Manuscript, Stuttgart University*.
- Duanmu, San (2007). *The phonology of standard Chinese*. Oxford University Press.
- Halle, Morris & Kenneth N Stevens (1997). The postalveolar fricatives of Polish. *Speech production and language: In honor of Osamu Fujimura* 13, 176–191.
- Jeng, Jing-Yi (2006). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies* 40:1, 27–48.
- Kallay, Jeffrey & Jeffrey Holliday (2012). Using spectral measures to differentiate Mandarin and Korean sibilant fricatives. *Thirteenth Annual Conference of the International Speech Communication Association*.
- Keating, Patricia A (1990). Phonetic representations in a generative grammar. *Journal of Phonetics* 18:3, 321–334.
- Kudela, Katarzyna (1968). Spectral analysis of Polish fricative consonants. *Speech Analysis and Synthesis* 1, 93–188.
- Ladefoged, Peter & Zongji Wu (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics* 12:3, 267–278.
- Lee-Kim, Sang-Im (2011). Spectral analysis of Mandarin Chinese sibilant fricatives. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1178–1181.

- Lee, Wai-Sum (1999). An articulatory and acoustical analysis of the syllable-initial sibilants and approximant in Beijing Mandarin. *Proceedings of the 14th International Congress of Phonetic Sciences*, vol. 413416.
- Li, F (2008). *The phonetic development of voiceless sibilant fricatives in English, Japanese, and Mandarin Chinese*. Ph.D. thesis, Ohio State University, Columbus.
- Li, Wen-Chao (1999). *A diachronically-motivated segmental phonology of Mandarin Chinese*, vol. 37. Peter Lang Publishing.
- Li, YS (2009). *Effects of Lexical Frequency and Neighborhood Density on Incomplete Neutralization in Taiwan Mandarin*. Ph.D. thesis, CCU.
- Lin, Yin-Hwei (2014). Segmental phonology. *The handbook of Chinese linguistics*, John Wiley & Sons, 400–422.
- McAuliffe, M, M Socolof, S Mihuc, M Wagner & M Sonderegger (2017). Montreal forced aligner [computer program] version 0.9.0, retrieved from <http://montrealcorpusools.github.io/montreal-forced-aligner/>.
- Nowak, Paweł M (2006). The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics* 34:2, 139–152.
- Padgett, Jaye & Marzena Żygis (2007). The evolution of sibilants in Polish and Russian. *Journal of Slavic Linguistics* 15:2, 291–324.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Solé, Maria-Josep (2007). Controlled and mechanical properties in speech. *Experimental approaches to phonology* 302–321.
- Stevens, Kenneth N, Zhiqiang Li, Chao-Yang Lee & Samuel Jay Keyser (2004). A note on Mandarin fricatives and enhancement. *From Traditional Phonology to Modern Speech Processing*, Foreign Language Teaching and Research Press, 393–403.
- Wickham, Hadley (2009). *ggplot2: Elegant graphics for data analysis*. Springer New York.
- Wolinski, Marcin, Marcin Milkowski, Maciej Ogrodniczuk & Adam Przepiórkowski (2012). Polimorf: a (not so) new open morphological dictionary for Polish. *LREC*, 860–864.
- Wu, Yuwen (1994). *Mandarin Segmental Phonology*. Ph.D. thesis, University of Toronto.
- Wu, Zongji & Maocan Lin (1989). *Overview of Experimental Phonetics*. Higher Education Press, Beijing.
- Żygis, Marzena & Silke Hamann (2003). Perceptual and acoustic cues of Polish coronal fricatives. *Proceedings of the 15th International Conference of Phonetic Sciences* 395–398.
- Żygis, Marzena, Daniel Pape & Bartłomiej Czaplicki (2012). Dynamics of sibilant systems: Standard Polish and its dialects. *Phonetik & Phonologie* 8, Jena: Friedrich-Schiller-Universität Jena.