Acoustic Comparison of Mandarin and Danish Postalveolars

Sidsel Rasmussen
Aarhus University

Mengzhu Yan
Victoria University of Wellington

Abstract
This paper examines similarities between the series of Mandarin "palatal" and "retroflex" affricates and fricatives. The distinct sets of Mandarin phones are known to be perceived as similar by nonnative listeners whose first language does not deploy the same postalveolar place contrast. Danish, for example, has phonological onset clusters articulated as alveolo-palatal sibilants, and previous studies indicate a tendency for Danish L1 speakers to confuse the two sets of Mandarin sounds. We obtain production data from native speakers of Danish and Mandarin and compare acoustic measurements to gain a better understanding of the cross-linguistic similarities and explore some of the perceptual problems indicated from perceptual assimilation data.

1. General introduction

It is well known that the sound inventories of the world's languages differ in how phonetic features distinguish contrastive categories, and a perceptual "retuning" may be one of the many tasks required of second language learners. Much of the existing literature on L2 acquisition research targets English sounds, but the body of literature targeting Mandarin Chinese is currently growing too, as more and more students around the world become interested in learning Chinese as a foreign language. The four lexical

Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen & Mette Hjortshøj Sørensen (Eds.), ***A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*** (pp. 65-80). Dept. of English, School of Communication & Culture, Aarhus University.

tones of Mandarin are notoriously difficult for learners of non-tonal languages to acquire, but what about acquisition at the segmental level? Mandarin Chinese has a rare three-way distinction of coronal sibilants: /ts, ts^h, s/, /tɕ, $tɕ^h$, ɕ/ and /tʂ, $tʂ^h$, ʂ/, in the literature typically referred to as the series of "dentals", "palatals" and "retroflexes", respectively. These are used as cover terms rather than precise denotations of place of articulation, and we follow the standard terminology throughout this paper. The "dental" sibilants have the most fronted constriction with the tongue against the incisors. The "palatals" are produced in the postalveolar region with the blade of the tongue, and they are often referred to as alveolopalatal consonants to specify a more fronted constriction than the term otherwise suggests. The "retroflexes" are apical post-alveolars (W.-S. Lee & Zee, 2003). In the following, we briefly introduce some of the research on Mandarin sibilant similarity before addressing the consequences of this crowded phonetic inventory for nonnative speech perception.

1.1. Introduction to Mandarin sibilant similarity

The Mandarin system of affricates and fricatives has been researched quite extensively. The palatal series is particularly interesting due to its seemingly predictable environment. By many analyses the palatals are said to appear only before the high front vowels [i] and [y] or the corresponding glides [j] and [ɥ], causing phonologists to treat them as allophones of other consonants (e.g. Hartman, 1944). A popular synchronic analysis posits them as allophones of the dentals (Duanmu, 2007), and previous studies have found the series of palatals to be acoustically most similar to the dental sibilants (S. Li & Gu, 2015; C.-Y. Lee, Zhang, & Li, 2014). However, the similarity (and perceptual confusability (Zhang, Lü, & Qi, 1982)) of this dental-palatal place contrast is presumably minimized by the nearly complementary distribution (Svantesson, 1986). Li and Zhang (2017) examined this claim experimentally and found that discrimination of the dental-palatal sibilant place contrast was less prone to errors when followed by [-high] nuclear vowels as opposed to an allophone of the high front vowel. Furthermore, they showed that the alternation of /i/ decreased the discrimination errors for both L1 and L2 listeners as well as increased listeners' response time as opposed to the condition in which both palatal and dental sibilants preceded [i]. The contrast between easily confusable Mandarin consonants seems to be enhanced by the alternation in the vowel quality as shown for the dental-palatal contrast, and the same is likely true for the retroflex series, which, just like the dentals, also combines with

a homorganic apical vowel instead of [i]. But how does discrimination fare in vowel contexts that do not offer an enhanced distinction which can facilitate consonant discrimination? For the non-high main vowels the distributional patterns between palatals and the other sibilant series are less distinct, if indeed present at all. For example, Ladefoged and Maddieson (1996) argue that the so-called "palatal" fricative followed by /a/ as the nuclear vowel does not have an intermediate glide linking the consonant and the low vowel: "from a phonetic point of view there is nothing other than a normal transition between the initial consonant and the following vowel" (p. 150). The consequence of this statement is a contrast neutralization which implies that a syllable like 下 <xia> ([ɕa]) forms perfect minimal pairs with e.g. [sa] and [ʂa], and this view resonates with recent phonological analyses that object to the notion of a medial glide in the traditional CGVX model[1] of the Mandarin syllable. Instead, a more complex consonantal onset inventory and a simpler CVX syllable structure has been proposed (e.g. Ao, 1992; Duanmu, 2017). The debate about Chinese medial glides is long and complicated, but if we look to these recent analyses and suspend the prescribed notion of the palatals always preceding high front segments, then the proposed complementary distribution of the traditional phonological analyses should not deter us from investigating consonants that are at least by some accounts, and in some environments, indeed minimally contrastive. This also means that the three-way place distinction in Mandarin sibilants may indeed be phonemic. Acoustic comparisons of these similar, yet contrastive sets of phones yield further insights into the exact cues listeners must attend to for successful discrimination.

1.2. The palatal-retroflex contrast in non-native perception

Norman (1988), among others, notes how L1 English learners of Mandarin tend to associate the palatals more closely with the retroflex series. This is quite interesting considering how the "palatal direction of confusability" is forward (towards the dentals) for native speakers, but backward (towards the retroflexes) for L1 English speaking learners. The L1 sibilant categories are very likely a source of influence here, given that English only has one place for articulating postalveolar sibilants, which might cause an "equivalence classification" of the distinct L2 phones. In this vein Chang

[1] In the traditional view the Mandarin syllable consists of maximally four segments of which only the nuclear vowel is obligatory: C (consonant) – G (glide) – V (vowel) – X (offglide or nasal coda).

et al. (2011) investigated the production of the Mandarin sibilants /s, ɕ, ʂ/ and the English /s, ʃ/ by L1 Mandarin speakers, native English speaking L2 learners of Mandarin, and heritage speakers of Mandarin who had grown up in English dominant communities. They found that the phones most likely to be merged in production across all groups were the Mandarin /ʂ/ and the English /ʃ/, but with heritage speakers distinguishing the five sibilants better than speakers in the two groups of late learners. Their findings point to an assimilation of the English and Mandarin postalveolars for both of the nonnative groups, which makes sense considering the similar articulatory descriptions of both fricatives. Danish, like English, has a contrast between alveolar and postalveolar sibilants: Specifically, Danish has a palatalized postalveolar sibilant [ɕ] produced either apically or laminally, typically denoted as an "alveolo-palatal" voiceless fricative, which is the phonetic realization of the onset sequence /s+j/ (Grønnum, 2009). Unlike English, however, Danish has no phonological affricates, but phonetically the clusters of /dj/ and /tj/ result in [tɕ] and [tɕʰ], respectively.[2] Just like Mandarin consonants, these Danish sibilant clusters are phonotactically restricted to onset position, and their phonetic similarity to the Mandarin alveolo-palatal sibilants is evident from the identical IPA symbols alone, but this type of comparison between transcriptions should only serve as an "initial heuristics in attempts to establish similarity" (Bohn, 2002: 198). Rasmussen and Bohn (2017) therefore examined the cross-linguistic mapping of Mandarin initial consonants as perceived by naive native Danish speaking listeners to test the perceived similarity between Mandarin and Danish consonants. 24 Danish L1 listeners with no prior knowledge of Chinese languages listened to Mandarin CV syllables over headphones and assimilated the initial consonant to a native onset represented in standard Danish orthography corresponding to unambiguous phonetic categories. The forced identification task was supplemented by a 7-point Likert scale on which, for each trial, participants indicated the perceived similarity between stimulus and the selected L1 response category. Results indicated that palatals and retroflexes were assimilated to the same native category of onset clusters with varying degrees of "goodness" of the perceived match. The theoretical implications of a two-to-one L2-L1 mapping is posited in the framework of the Perceptual Assimilation Model (Best, 1995).

[2] We refer to the Danish sibilants by their phonemic representations throughout this paper in order to avoid confusion with the Mandarin alveolopalatals transcribed by the same symbols. We find it useful to deploy the impressionistic labels "lenis" and "fortis" when referring to the collective groups of unaspirated and aspirated affricates, respectively.

If two nonnative sounds are heard as equally good variants of a single native phoneme, discrimination difficulties are predicted. This is known as a Single Category assimilation type (SC). If the nonnative phones are assimilated to the same native category with *differing* ratings of the match, a Category-Goodness (CG) difference is observed, and discrimination between the L2 segments should be less problematic. Findings from the perceptual assimilation study revealed a strong effect of the following vowel on the cross-linguistic matching of Mandarin palatals. Considering the phonotactics of Mandarin as well as the restrictions imposed by the alternating vowel quality on Mandarin syllables, the data included here only lists results for the vowel context for /a/, which is preceded by both retroflexes and palatals.

DA Response	MA Stimuli					
	/tɕ/	**/tɕʰ/**	**/ɕ/**	**/tʂ/**	**/tʂʰ/**	**/ʂ/**
/dj/	**47 (5.53)**			**60 (5.14)**		
/tj/	35 (4.08)	**78 (5.16)**		22 (5.25)	**71 (3.98)**	
/sj/			**74 (4.11)**			**85 (4.85)**

Table 1. Confusion matrix of Mandarin (MA) stimuli presented before /a/ (horizontally) and selected Danish (DA) response (vertically). First number indicates percentage of match, and number in parentheses is average "goodness". Bolded cells show most frequent match.

As Table 1 shows, the palatal-retroflex place contrast is not attended to by naïve Danish listeners, who identify both series of Mandarin tokens as above average exemplars of their articulatorily closest native counterpart. Interestingly, there does not seem to be a clear preference for Mandarin palatals which as the most similar to the Danish sibilants if we are to trust phonetic descriptions: only the palatal aspirated affricate fares better than its retroflex counterpart in both number of matches and similarity rating. The unaspirated Mandarin affricates are most frequently perceived as fairly good exemplars of Danish /dj/, but surprisingly, the fortis /tj/ response is also selected as the preferred match for 22-35% of trials. The results from this small part of a perceptual assimilation study yield two main questions of interest: Firstly, if Danish sibilants are produced as the Mandarin

palatals, why is there for unaspirated Mandarin sibilants seemingly a (small) preference for the retroflexes in a cross-linguistic mapping task? Secondly, why would the Mandarin unaspirated affricates sometimes be perceived as good exemplars of Danish /tj/? To shed light on the differences and similarities between the nine sibilants discussed here it will be useful to examine spectral properties of the contrasts. Only one study has so far examined the Danish-Mandarin sibilant contrast acoustically. Mikkelsen (2016) studied Danish L2 Mandarin learners' production of the L1 [ɕ], the L2 English [ʃ] and L3 Mandarin [ɕ] and [ʂ]. She measured COG[3] values for the four fricatives produced in different vowel contexts and found that the Mandarin postalveolars were produced more similarly preceding non-high vowels. This can probably be explained as an effect of the differences in vowel quality on the preceding consonant. In the /a/ condition, the retroflex and the Danish fricative were not produced significantly differently, while the other sibilants were statistically distinguished by the spectral means obtained. Mikkelsen's data, although limited to few tokens per target consonant, indicate that Danish learners of Chinese produce the retroflex fricative similar to their native category /sj/. It also indicates that there is an effect of following vowel, a finding in agreement with observations from other acoustic studies of native Mandarin production as well as the perceptual assimilation study mentioned above. Additional acoustic measures, a larger data set and inclusion of the homorganic affricates will likely provide additional insights to these previous cross-linguistic studies.

1.3. The current study

In this paper we revisit the two-to-one mapping of Mandarin palatals and retroflex consonants to Danish onset clusters /dj, tj, sj/. As Bohn (2002) notes, direct comparison of cross-linguistic similarity is often not possible due to methodological differences between the studies that report acoustic measurements. The most informative and best comparable production data will require attention to factors such as elicitation method and phonetic environment of target sounds. In this study we present data from both languages, elicited specifically for the purpose of allowing for comparisons between relevant acoustic measures of the two languages in order to explore an additional aspect of the phonetic similarity indicated in a previous study. We compare our Mandarin data to existing literature on the acoustics of Mandarin sibilants. We are not aware of any acoustic studies examining the

[3] Center of gravity is the spectral mean, often used to classify aperiodic speech sounds such as fricatives.

Danish target categories under investigation, so this study can hopefully both serve the purpose of detailed acoustic comparisons between Mandarin and Danish postalveolars, probing the questions raised from a perceptual assimilation study, as well as provide baseline acoustic measures for three Danish consonant clusters.

2. Methods

2.1. Participants

Five female native speakers of each language were recorded. The five native Mandarin speakers (mean age: 27.6, SD: 3.0) were recorded in New Zealand, but all reported frequent use of L1 Mandarin Chinese in their everyday communication. The Danish speakers had a mean age of 26.6 (SD: 3.1) and were recorded in Aarhus, Denmark. A basic language background questionnaire ensured that they had no prior experience with any Chinese languages. All ten speakers participated as volunteers in this study, and none of them reported any speech or hearing problems.

2.2. Recordings

Four Mandarin affricates /tʂ, tʂʰ, tɕ, tɕʰ/ and two fricatives /ʂ, ɕ/ were included in the Mandarin stimuli list. The target consonants were combined with /a/ and the falling tone (T4). Each target phoneme was repeated ten times, resulting in a total of 300 tokens (6 target consonants * 1 vowel * 1 tone * 10 repetitions * 5 speakers). The syllables were embedded in the middle position of a carrier sentence 我把__ 读出来 ("I take __ and read out loud.") and the corresponding Pinyin Romanization and tone number were also written after the target syllable. The Danish stimuli were created with the intent to provide a context as similar to the Mandarin syllables as possible. Unlike the Chinese stimuli, the Danish target syllables do not correspond to morphemes, so elicitation relied on speakers' production of nonsense syllables for which rhyming real words were provided prior to recording. To achieve similarity in vowel context we created a list of open syllable non-words that combined the target Danish onsets with a low back vowel [ɑ]. The back vowel allophone surfaces due to a fusion between /a/ and /r/ (Basbøll, 2005: 149), so to ensure the anticipated vowel quality in the targets, the orthographic representations listed were <*djar*>, <*tjar*>, <*sjar*>. High frequency words (e.g. '*har*' [hɑˀ][4] (present tense of "to

[4] We anticipated glottal stop (stød) on the open monosyllables in focus position. Mandarin T4 with a falling contour has the shortest duration (Ho, 1976) and T4 syllables are perceptually not unlike syllables with stød.

have")) were provided as rhymes to targets. All Danish targets and fillers were produced in the carrier sentence *Jeg siger __ til dig*. ("I say __ to you."). Both sets of stimuli were pseudo randomized so that no more than two identical sentences occurred together. Non-target fillers were included as the last sentence in the written material to avoid end of list intonation. For the longer set of stimuli (the Mandarin list), breaks were provided. Each session lasted approximately 1-2 minutes.

2.3. Measurements

The segmentation and measurements of the target fricatives and affricates were handled separately by the authors in Praat (Boersma & David Weenink, 2018). The task was divided between the authors so the stimuli matched the L1 of the author in order to also perceptually verify the onset labels, which resulted in the discarding of 1 Mandarin and 4 Danish tokens. Speech segmentation was conducted on the basis of waveform and wideband spectrogram. For fricatives, the entire noise portion was classified as the target consonant. For affricates, frication noise was defined between the beginning of the burst and the onset of the vowel. A number of studies have provided acoustic analyses for Mandarin, and different acoustic measurements have been found to differentiate the Mandarin sibilants. For the place distinction, spectral moments (notably the spectral mean, i.e. COG) and F2 frequency at onset of following vowel have been good discriminators (C.-Y. Lee et al., 2014; S.-I. Lee, 2011; S. Li & Gu, 2015). Manner differences, such as state of aspiration in the affricates has been shown to be distinguished well by duration and amplitude (S. Li & Gu, 2015). For this study, we extracted the following five acoustic parameters using Praat: normalized duration of frication and the four spectral moments: center of gravity (COG), dispersion, skewness and kurtosis. ProsodyPro (Xu, 2013) was used to generate the actual durations of segment proportions, and following S. Li & Gu (2015) we calculated the normalized durations (i.e. the ratio of the consonant duration to the duration of the entire syllable) to avoid the influence of speaking rate. A script (Mayer, 2011) was used for obtaining measures for the four spectral moments calculated over the middle 40 ms of the frication portion.

3. Results

A set of linear regression models were used to analyze the data (i.e. normalized duration, four spectral moments) as the dependent variable and interaction between place of articulation and manner as independent

variables using R (R Core Team, 2018). In order to see which groups differed from each other, 'esmeans' (Lenth, Singmann, Love, Buerkner, & Herve, 2018) was used for pairwise comparison.

3.1. Normalized duration of frication

Table 2 lists the values for the nine target consonants, the entire target syllable as well the duration proportion of the consonant relative to the syllable. All values are averaged across speakers and repetitions. The Danish alveolo-palatals are termed "Palatal" for short in the following presentations of data. Data is arranged by the three sets of sibilant (Danish, Mandarin palatals and Mandarin retroflexes, i.e. a presumed place contrast), and by manner (fricative = blue, unaspirated affricate = green, aspirated affricate = red).

Place	Onset consonant	Consonant duration (ms)	Syllable duration (ms)	Normalized duration*
Danish Palatal	/dj/	57.01	347.20	0.16
	/tj/	130.96	390.17	0.34
	/sj/	181.30	420.50	0.43
Mandarin Palatal	/tɕ/	72.26	289.25	0.25
	/tɕʰ/	130.57	344.22	0.38
	/ɕ/	153.04	432.17	0.35
Mandarin Retroflex	/tʂ/	50.57	274.60	0.18
	/tʂʰ/	117.85	324.43	0.36
	/ʂ/	153.38	400.41	0.38

Table 2. Consonant duration, syllable duration and normalized duration. *Normalized duration is calculated as a mean of all the individual token's normalized duration rather than as a ratio of the two means given in the previous columns in the chart.

Table 2 and Figure 1 show that the manner of articulation (i.e. aspirated affricates, unaspirated affricates and fricatives) played an important role in the duration of the target consonants, such that fricatives were the longest followed by aspirated affricates and unaspirated affricates. The role of the

aspiration was in line with measurements presented by S. Li & Gu (2015), who found the same hierarchy of frication duration between the three manner distinctions. As Figure 1 shows, the Danish short lag affricate /dj/ did not differ from /tʂ/ in terms of normalized duration (t=-1.061, p=0.9793), but /dj/ was shorter than /tɕ/ (t=4.104, p=0.0016). Duration of the Danish long lag (and aspirated) affricate /tj/ was similar to /tʂʰ/ and /tɕʰ/ (/tj/ vs. /tʂʰ/: t = -1.317, p=0.9260; /tj/ vs. /tɕʰ/: t=-2.570, p=0.2022). Duration of /sj/ was similar to both /ʂ/ and /ɕ/ (/sj/ vs. /ɕ/: t=-0.720, p=0.9985; /sj/ vs. /ʂ/: t=-0.648, p=0.9993). Duration proportions indicate that the Danish fricative and aspirated affricate are similar to both of their Mandarin counterparts while Danish unaspirated /dj/ is produced most similar to retroflex /tʂ/, and is clearly produced with a shorter frication proportion than Mandarin /tɕ/.

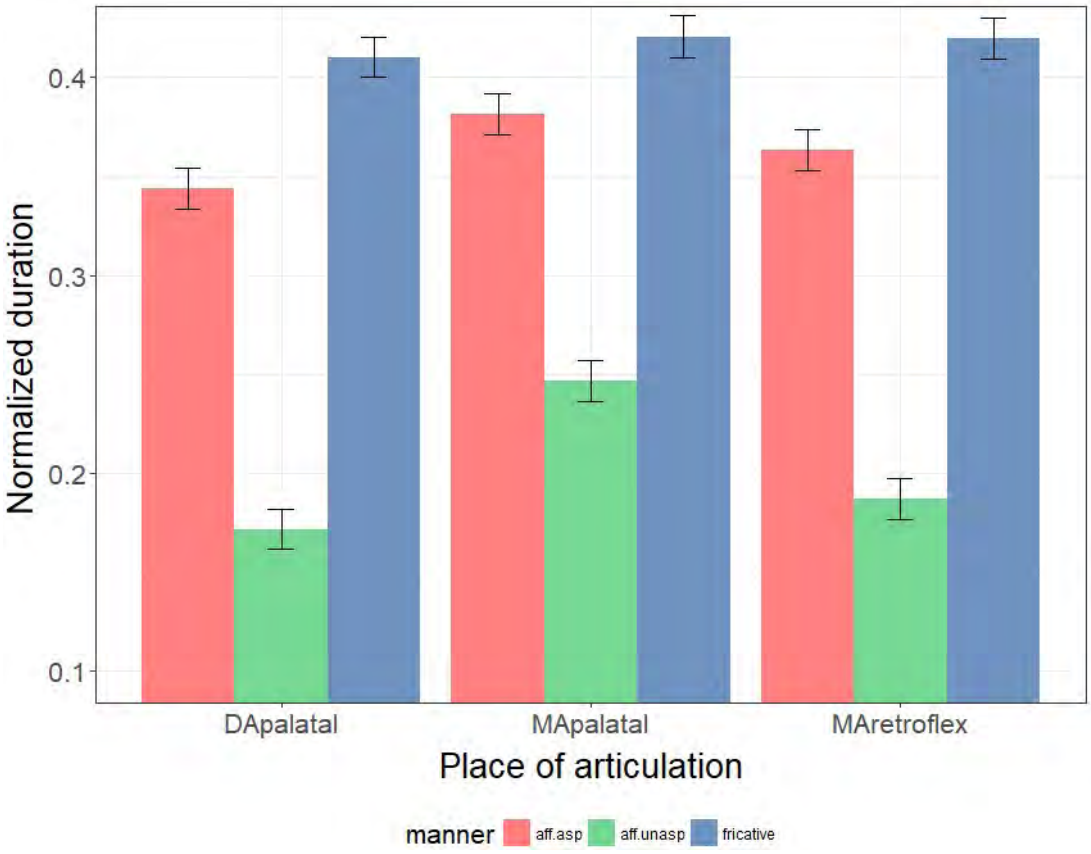

Figure 1. Normalized durations (proportional duration of C relative to the entire syllable). Error bars indicate standard errors.

3.2. Spectral moments

Table 3 displays values for four spectral moments averaged across speakers and repetitions.

Place	Onset consonants	COG (Hz)	Dispersion (Hz)	Skewness	Kurtosis
Danish Palatal	/dj/	5638	2042	2.28	8.56
	/tj/	5714	2123	2.08	6.53
	/sj/	5571	2112	2.23	7.90
Mandarin Palatal	/tɕ/	8284	1864	1.16	3.65
	/tɕʰ/	7714	2006	1.30	3.17
	/ɕ/	8103	1916	1.30	3.37
Mandarin Retroflex	/tʂ/	5737	2074	1.27	2.70
	/tʂʰ/	5496	2368	1.23	2.39
	/ʂ/	5652	2278	1.26	2.35

Table 3. Mean spectral moments.

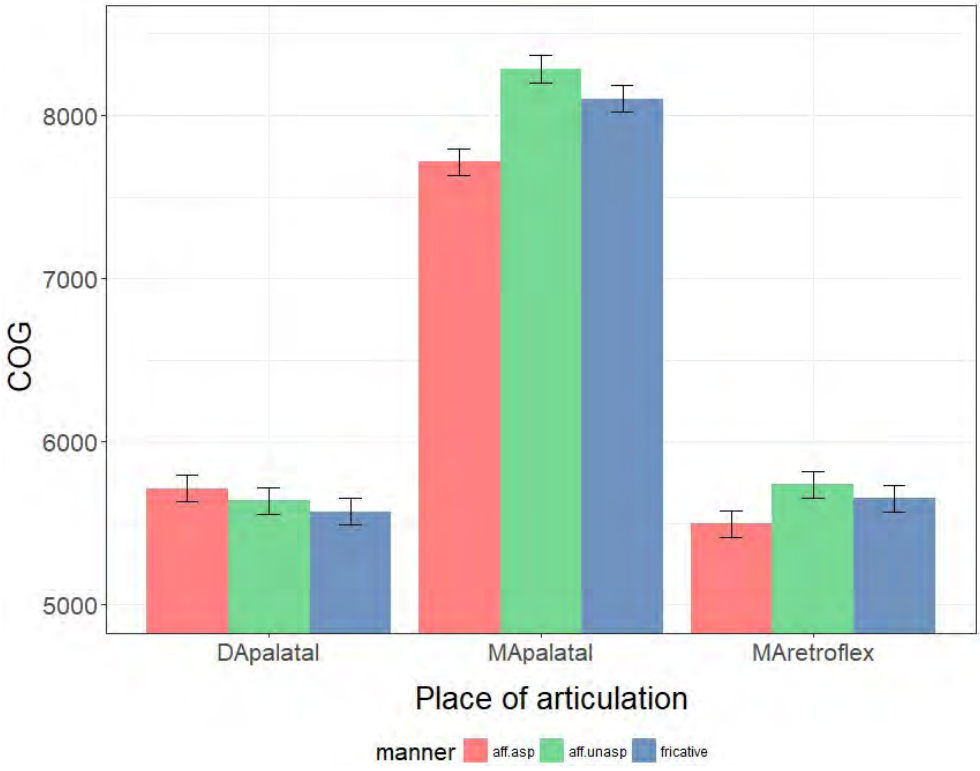

Figure 2. Center of gravity.

As shown in Table 3 and Figure 2 above, place of articulation was the main effect on COG. Post-hoc pairwise comparisons revealed that for Mandarin onsets, palatals had higher COG than retroflexes (p<0.05), in line with findings from previous studies (e.g. S.-I. Lee, 2011) and consistent with the fact that higher COG values are found for fricatives produced anteriorly (e.g. Jongman, Wayland, & Wong, 2000). The Danish alveolo-palatals were not significantly different from Mandarin retroflexes (p>0.05) which is in line with what Mikkelsen (2016) found in her cross-linguistic comparison of fricatives. In terms of the second spectral moment, dispersion, Danish palatals were not significantly different from their corresponding Mandarin palatals (p>0.05). However, Danish aspirated palatal /tj/ had a smaller standard deviation compared to the Mandarin palatal /tɕʰ/ (t=-3.515, p=0.05) resulting from less variability among the Danish aspirated affricate tokens than their Mandarin palatal counterparts. Danish fricative /sj/ had similar standard deviation as Mandarin /ɕ/ (t=2.877, p=0.0972). Regarding the third and the fourth spectral moments, Danish sibilants differed significantly from all Mandarin onsets (p<0.05). Svantesson (1986) plots the two first spectral moments and displays how the fricative tokens from his four male speakers form clusters that are not clearly separate from one another. We similarly graph dispersion as a function of COG in Figure 3 and Figure 4 below to visually represent the overlaps in acoustic space for three sibilants *within* each series, as well as *between* the Danish "palatal" and the Mandarin retroflex series (compare Figures 3 and 4).

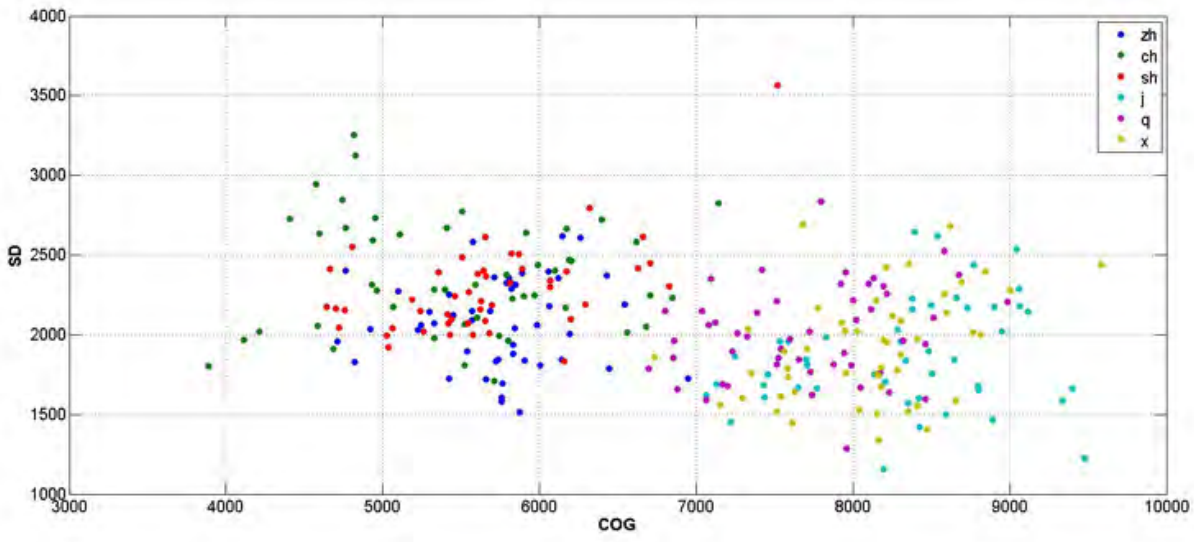

Figure 3. Mandarin sibilants. Items are labelled in pinyin in the legend: The retroflex series /tʂ, tʂʰ, ʂ/ corresponds to <zh, ch, sh> and the palatal series / tɕ, tɕʰ, ɕ/ corresponds to <j, q, x>.

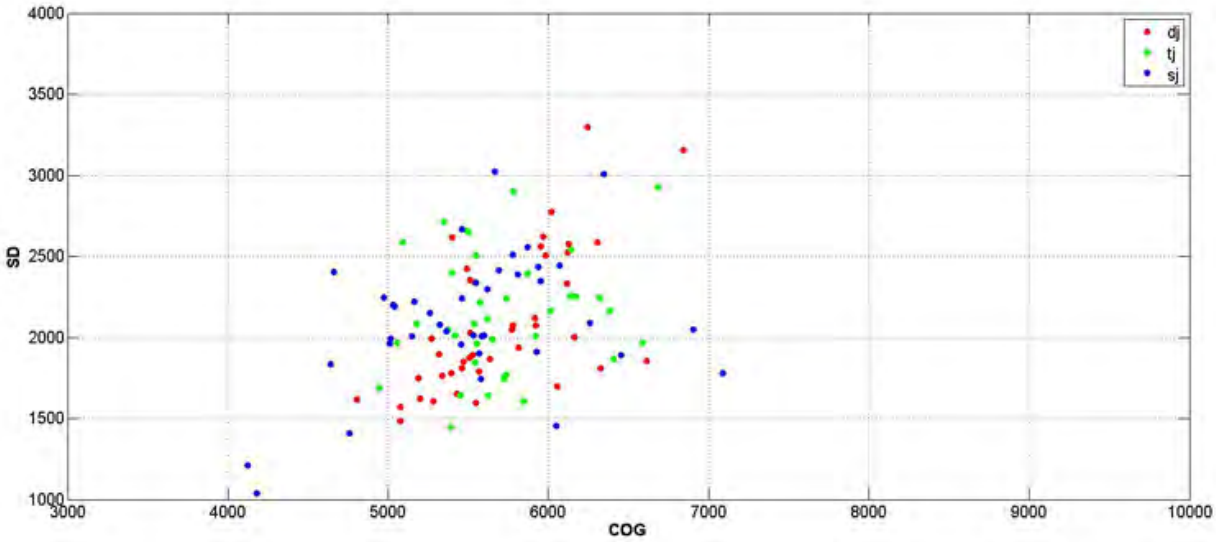

Figure 4. Danish sibilants.

4. Discussion

This study addressed some of the questions arisen from a cross-linguistic assimilation study, which investigated the mapping of Mandarin consonants by naïve Danish listeners. The previous study indicated that, preceding low vowels, two sets of Mandarin coronal sibilants /tɕ, tɕ^h, ɕ/ and /tʂ, tʂ^h, ʂ/ were assimilated to only one set of Danish sibilants /dj, tj, sj/, with Mandarin retroflexes typically faring slightly better than the palatals in terms of assimilation percentage. This modest preference for retroflexes as matches of the native sibilants was surprising considering the respective articulatory descriptions of categories, from which a closer similarity with the Mandarin so-called "palatal" series would be expected. Our results from the current study, however, reveal that the Danish sibilants are acoustically more similar to the Mandarin retroflexes than the palatals, most clearly shown in the comparison of spectral moments. COG values for the series of Danish sibilants were significantly different from values obtained for the Mandarin palatal consonants, while measures for the Danish onsets and the Mandarin retroflexes were not significantly different. Since Danish only has one series of comparable postalveolar sibilants, the two-to-one mapping is nevertheless quite expected. The acoustic data presented here would suggest that the assimilation of Mandarin postalveolars preceding /a/ should yield a Category-Goodness distinction, with the retroflex series now identified as the most similar in terms of the acoustic signal. As COG measures do not reliably distinguish Danish sibilant onsets from Mandarin retroflexes in the production of two groups of native speakers, Mikkelsen's

(2016) production data from Danish learners of Mandarin does not necessarily indicate the merging of a non-native and a native category.

The second question examined in this study concerned the unexpected assimilation of Chinese unaspirated affricates to the Danish aspirated /tj/. While /dj/ was still the preferred response, the native aspirated affricate /tj/ was sometimes selected as the closest native match to both /tʂ/ (22%) and /tɕ/ (35%). Since duration is one of the cues known to distinguish a [+/-aspiration] contrast we compared the normalized durations of the three sets of unaspirated and aspirated affricates. Our results show that the acoustic properties of Danish /dj/ differ significantly from all Mandarin affricates except /tʂ/, and /tj/ as significantly different from both /tʂ/ and /tɕ/. The normalized durations therefore do not offer any direct explanation for the unexpected mapping, and additional measurements might be needed to be able to fully account for the unexpected perception results. We speculate, however, if the duration differences between the three lenis onsets might hint at an explanation: The fact that the Mandarin affricate /tʂ/ is minimally longer and /tɕ/ is significantly longer than /dj/ might mean that duration for both of the Mandarin unaspirated affricates actually exceeds the perceptual boundary for the Danish /dj/ category, causing native listeners to classify them as acceptable variants of their native /tj/ instead.

Our data also brings into question the articulatory descriptions of these Danish consonant clusters, indicating a more retracted point of constriction in the oral cavity than what is typically assumed. It is, however, still possible that a combination of Danish sibilants + high vowel would yield a more fronted articulation of the consonant clusters, thereby increasing the spectral mean, resulting in an acoustic signal resembling of that for Mandarin palatals. Future studies should investigate how different sibilant + vowel combinations might account for differences in the spectral cues that discriminate sibilant categories. The discrepancy between articulatory descriptions and this newly obtained acoustic data of Danish sibilants is also well worth exploring in further detail.

Acknowledgements

We are grateful to our ten speakers and to Jonas Villumsen for obtaining the Danish recordings while neither of us were in Denmark and to Allard Jongman for his helpful comments on an earlier version of the paper. We are likewise very grateful to the reviewer for helpful comments. Any errors are our own. And finally, we are incredibly grateful to Ocke for encouraging us to start working in the field of phonetics, for being an inspiring teacher and a supportive advisor.

References

Ao, B. (1992). Non-uniqueness condition and the segmentation of the Chinese syllable. In E. Hume (Ed.), *Ohio State University: Working papers in linguistics* (pp. 1-15). Columbus: Ohio State University.

Basbøll, H. (2005). *The phonology of Danish*. Oxford / New York: Oxford University Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore, MD: York Press.

Boersma, P., & David Weenink. (2018). *Praat: Doing phonetics by computer (Version 6.0.39)*. Retrieved from www.praat.org

Bohn, O.-S. (2002). On phonetic similarity. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 191-216). Trier, Germany: Wissenschaftlicher Verlag.

Chang, C. B., Haynes, E. F., Yao, Y., & Rhodes, R. (2009). A Tale of Five Fricatives: Consonantal Contrast in Heritage Speakers of Mandarin. *University of Pennsylvania Working Papers in Linguistics, 15*(1), 36-43.

Duanmu, S. (2007). *The phonology of standard Chinese* (2nd ed). Oxford / New York: Oxford University Press.

Duanmu, S. (2017). From non-uniqueness to the best solution in phonemic analysis: Evidence from Chengdu Chinese. *Lingua Sinica, 3*(1). https://doi.org/10.1186/s40655-017-0030-7

Grønnum, N. (2009). *Fonetik og fonologi: Almen og dansk* (3rd ed.). København: Akademisk.

Hartman, L. M. (1944). The segmental phonemes of the Peiping dialect. *Language, 20*, 28-42.

Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica, 33*(5), 353-367. https://doi.org/10.1159/000259792

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*(3), 1252. https://doi.org/10.1121/1.1288413

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK / Cambridge, Mass., USA: Blackwell Publishers.

Lee, C.-Y., Zhang, Y., & Li, X. (2014). Acoustic characteristics of voiceless fricatives in Mandarin Chinese. *Journal of Chinese Linguistics, 42*(1), 150-171.

Lee, S.-I. (2011). Spectral analysis of Mandarin Chinese sibilant fricatives. *18th ICPhS, Glasgow*, 1178-1181.

Lee, W.-S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association, 33*(1), 109-112. https://doi.org/10.1017/S0025100303001208

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Estimated Marginal Means (Version 1.2.3) [R]. Retrieved from https://github.com/rvlenth/emmeans

Li, M., & Zhang, J. (2017). Perceptual distinctiveness between dental and palatal sibilants in different vowel contexts and its implications for phonological contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8*(1), 1-27. https://doi.org/10.5334/labphon.27

Li, S., & Gu, W. (2015). Acoustic analysis of Mandarin affricates. *Sixteenth Annual Conference of the International Speech Communication Association / INTERSPEECH 2015, Dresden.*

Mayer, J. (2011). *Calculate spectral moments*. Retrieved from http://praatpfanne.lingphon.net/downloads/spectral_moments.txt

Mikkelsen, S. (2016). *Chanish? A study of a Danish accent in the production of two Chinese sibilants* (Unpublished master's thesis). Aarhus University, Aarhus.

Norman, J. (1988). Chinese. Cambridge / New York: Cambridge University Press.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Rasmussen, S., & Bohn, O.-S. (2017). Perceptual assimilation of Mandarin Chinese consonants by native Danish listeners. *The Journal of the Acoustical Society of America, 141*(5), 3518-3518. https://doi.org/10.1121/1.4987399

Svantesson, J.-O. (1986). Acoustic analysis of Chinese fricatives and affricates. *Working Papers in Linguistics, 25*(1), 195-211.

Xu, Y. (2013). ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody*. Aix-en-Provence, France.

Zhang, J., Lü, S., & Qi, S. (1982). A cluster analysis of the perceptual features of Chinese speech sounds. *Journal of Chinese Linguistics, 10*(2), 190-206.