Contents

Abstract

Baseret på modeller for indlæring af fremmedsprogsfonologi, forudsagde Wang (2014) at indlæringen af de kinesiske dentale obstruenter ville være særligt problematiske at lære for danskere. Denne antagelse skyldtes at kinesisk både har fonemiske dentale plosiver og affrikater /t tʰ ts tsʰ/, og at aspirationen i den danske aspirerede alveolære plosiv manifesteres som affrikation /tˢ/. Til gengæld er det sidenhen blevet vist at danskere med blot kortvarig erfaring med kinesisk uproblematisk kan opfatte forskellen mellem de kinesiske dentale obstruenter, og at selv danskere uden erfaring med kinesisk klarer det overraskende godt. I denne opgave testes det hvorvidt produktionen af de pågældende lyde er ligeså uproblematisk. Desuden testes en hypotese om at danskeres /tˢ/ ofte overføres til deres udtale af /tʰ/ på engelsk, hvilket yderligere ville forventes at problematisere indlæringen af det tilsvarende fonem på kinesisk. Dette testes gennem analyse af optagelser af 25 studerende af kinesisk på forskellige niveauer, der indtaler korte sætninger på dansk, engelsk og kinesisk.

Resultaterne viser utvetydigt at deltagerne i eksperimentet overfører affrikation til deres udtale af engelsk, og at mange danskere der lærer kinesisk har problemer med produktionen af de dentale obstruenter. Mens /ts/ tydeligt udtales distinkt fra /tʰ tsʰ/, udtales den ofte blot som en frikativ, og varigheden er markant længere end hos førstesprogstalere af kinesisk. Hvad angår /tʰ tsʰ/, tyder det på at nogle kinesisk-studerende opfatter variation i varighed og ikke affrikation som den primære forskel på de to. I hvert fald er affrikation prominent i udtalen af begge fonemer for de fleste studerende både på første og andet år, og blandt halvdelen af de deltagende studerende på tredje år. For at komplicere dette yderligere, udvikler studerende løbende en fortis og dermed længere udtale af kinesisk /tʰ/. Dette er uproblematisk hvis deres udtale er rent aspireret, men hvis de har affrikation i udtalen, kommer /tʰ tsʰ/ til at udtales ens i endnu højere grad.

The productive acquisition of dental obstruents by Danish learners of Chinese

- RASMUS PUGGAARD[1]

1 Introduction

The interaction between two phonologies that occurs in L2 acquisition is interesting no matter which two languages are being studied, because of the minor phonetic differences between otherwise similar phonemes in the two languages. In a paper presented at Fudan University in October 2014 (Wang 2014), some specific production and perception difficulties were predicted for Danish speakers learning Standard Chinese (SC) as a second language (L2). These predictions were based on existing theoretical models for the acquisition of L2 phonology, and specifically the acquisition of SC dental obstruents was emphasized. There is plenty of literature on L2 acquisition of SC phonology, particularly on the acquisition of tones (e.g Yang 2015). But the study of SC acquisition by native speakers of Danish is a new field, albeit one that has been very active lately. Recent work on Danes' acquisition of SC includes a study by Sloos et al (2015) on the acquisition of speech rhythm, finding that native speakers of Danish do not observe SC intrinsic tone duration, and do not have the characteristic final syllable lengthening of SC; Sloos et al (forthc. a; forthc. b) tested the pitch range and tone realization of native Danes' SC, and found that their pitch variability was significantly lower than that of native speakers, and that tone realization was often fairly poor, presumably due to the very small pitch range of Danish; Ne et al. (forthc.) tested Danish learners' perception of the SC dental stops and affricates, and surprisingly found that 1st year students of Chinese showed near-native discrimination of them, and that even Danes with no knowledge of Chinese were mostly able to distinguish between them. But as reported by a teacher of Chinese at Aarhus University, productive acquisition of the sounds appears to be problematic[2]. By comparing the phonologies of Danish and SC, this paper sets up a hypothesis for what problems Danish speakers may have when producing these sounds, and tests it by analyzing the production of the dental stops and affricates by Danish learners of SC through recordings made of students at different levels of their studies. As predicted by Wang (2014), it was found that a specific quirk of Danish phonology, namely the affrication of alveolar stops, results in difficulties with the production of the SC dental obstruents. The problem is tenacious, and is still found for some 3rd year students.

In Section 2 below, three models accounting for the perception of L2 sounds will be presented. Section 3 analyzes the phonology of Danish and SC – specifically the simple onset consonants – and makes predictions for Danish learners' implementation of SC consonants using the models presented in Section 2. Section 4 will present the speakers, stimuli and procedure used in the recordings, before Section 5 gives the method used for analyzing the recordings. In Section 6 and Section 7 the results of two experiments are analyzed, the first being a small experiment

[1] Parts of this paper have been previously presented at the International Symposium on Chinese Theoretical and Applied Linguistics at Newcastle University, December 10, 2015.
[2] Zhang Chun, personal communication

testing Danes' pronunciation of aspirated alveolar stops in English. The results are discussed in relation to the predictions in Section 8, while Section 9 summarizes the paper.

2 Theoretical background

Developing productive competence of the phonological system of an L2 as an adult is not easy. Flege (1980:119) notes that the development of a new phonological system appears slow compared to more abstract linguistic competences such as development of lexicon and syntax. When learning an L2 as an adult, the native language (L1) phonological system interferes with that of the L2 in complex ways, both in perception and production of the L2. There are three well-known models accounting for how L2 sounds are filtered through the phonology of the L1: Kuhl's *Native Language Magnet* model (NLM; e.g Kuhl et al 1992); Best's *Perceptual Assimilation Model* (PAM; e.g Best et al 2001); and Flege's *Speech Learning Model* (SLM, e.g Flege 1995).

The NLM proposes that listeners develop acoustic prototypes for L1 phonemes very early in life. These acoustic prototypes serve as "magnets" in the surrounding perceptual space, making it harder to perceive variation surrounding L1 prototypes. Different nonprototypic members of a category will be perceived as more similar to the prototype than to each other, regardless of the actual physical difference between them (Kuhl et al 1992:607). Likewise, according to the NLM, L2 phones that are similar to L1 prototypes will be dragged into the perceptual space of the prototypes, and listeners will not develop prototypes for novel L2 categories (Best et al 2001:776). Lively and Pisoni (1997) report that the results supporting the NLM are much more robust for small children than for adults, and also find that phoneme prototypes are not stable across different phonetic contexts.

The PAM was originally developed to account for the fact that speakers of American English were found to discriminate Zulu clicks surprisingly well; according to PAM, this is because the clicks were not perceived as speech sounds, in which case L1 phonology neither aids nor hinders discrimination (Best et al 1988). L2 sounds that are not similar to any L1 sounds are deemed Non-Assimilable speech sounds, and discrimination among them is expected to be as good as the actual difference between the acoustic cues allows. Given that there is no interference from L1 phonology in these cases, discrimination is expected to range from good to excellent. Discrimination between two L2 sounds is also expected to be accurate when they are separated by phonological boundaries that also exist in L1, or if one is similar to an L1 phoneme while the other cannot be easily identified with a specific L1 phoneme. When two L2 sounds are equidistant to one L1 phoneme, however, discrimination is expected to be poor. This is referred to as Single Category assimilation. Two L2 sounds may also be similar to one L1 phoneme, but one is more similar than the other; this is termed a Category Goodness difference. Here, discrimination is also expected to be poor, but not as poor as in Single Category discrimination (Best et al 2001:777).

Of the three models, the SLM is the only one with primary focus on production of L2 sounds as opposed to perception, although the two are assumed to be closely linked. The model also mainly focuses on the production of fairly advanced L2 learners. The SLM assumes that L2 sounds are classified on the basis of perceptive similarity to the nearest native sound. The further the distance is from an L2 sound to the nearest L1 sound, the better the chance is for a new

stable sound category to develop. If an L2 sound is reasonably close to the nearest L1 sound, there is a good chance of it being perceived as part of the L1 category, which may result in production like a good exemplar of the L1 category, even though that may be a poor exemplar of the L2 sound (Flege 1995:239). If any phonetic differences are discerned by the learner, a new phonetic category may be developed; however, the learner's category may not be similar to the corresponding category of a native speaker, since it may be based on different features. An example of this is the production of English stops by native speakers of Canadian French: the English aspirated stops /p t k/ have a much longer release duration than the corresponding French voiceless stops. Caramazza et al (1973, cf. Flege 1981:450) report that L1 speakers of French produce English stops with a release duration somewhere in between their own French and that of an L1 speaker of English, indicating that while they have different categories for English and French /p/, their English /p/ category is still different from that of a native speaker. Besides, the SLM maintains that some L2 production errors are not perceptually motivated, but due to motoric output constraints from the L1 (Flege 1995:238).

Apart from the interference of one phonological system with another, there is also ample evidence that there is an influence of L1 orthography-phonology correspondences when learning an L2. This explanation should only be used when an observed problem cannot be explained using a phonological model; for example, Bassetti (2006a, cited in Bassetti 2008:197) found that Italian learners of Chinese were likely to pronounce Chinese voiceless unaspirated stops /p t k/ as voiced /b d g/ due to Pinyin's use of <b d g> for these. Meng (1998, cited in Bassetti 2006b:100) made a similar case for English learners of SC. It is my view that this problem can – and should – just as easily be explained using the phonological models presented above.

3 Phonology of Danish and SC

There are many challenges for the Danish learner wishing to learn SC as an L2; a complex tone system has to be developed from scratch; in spite of the huge vowel phoneme inventory of Danish, there are still new ones to be learned; the learner must learn to use retroflex consonants; the learner must master a prohibitively complex new writing system. These challenges are not unique to the Danish learner, but there is one aspect of the Danish phonology that poses a unique challenge to Danes wishing to learn the language: the Danish alveolar stop, which is "affricated" in simple onsets.

Given that the SC syllable only allows simple onsets (and a very restricted coda of either /n ŋ ɻ/), and given Flege's hypothesis that L1 and L2 sounds are related to each other at a position-sensitive allophone level and not a phoneme level (1995:239), only the simple onset allophones of Danish will be taken into account here. In Section 5.1 and 5.2 below, consonants allowed in simple onset of Danish and SC will be presented. In Section 5.3 the mapping possibilities of the phonemes will be analyzed, highlighting the problem of the dental obstruents.

3.1 Danish

3.1.1 Danish consonants in onset position

Following Grønnum (1998:39) the initial consonant allophones of Danish can be seen in Table 1.

Table 1: Initial consonant allophones of Danish

	Labial	Alveolar	Palatal	Velar	Uvular	Glottal
Unaspirated plosive	b̥	d̥		g̊		
Aspirated plosive	pʰ	tˢ		kʰ		
Fricative	f	s				h
Nasal	m	n				
Approximant	ʊ		j		ʁ	
Lateral		l				

Grønnum uses the less descriptive /b d g p t k/ for the stop phonemes because their realizations differ in coda position, but in the context of this paper I prefer to use onset-specific transcriptions. The exact position and manner of articulation of the Danish rhotic is hard to describe, as it is not static at any point and has elements of both a glide and a trill (Grønnum 2007:115).

Unlike e.g English, Danish stops do not have a voicing distinction, but are exclusively voiceless. Neither do Danish stops show a difference in fortis and lenis, but are exclusively lenis (Grønnum 1998:107, 263). This gives the set /b̥ d̥ g̊ pʰ tˢ kʰ/, though narrowly the simple onset allophones of the aspirated set can be given as [b̥ʰ, d̥ˢʰ ~ d̥ˢ[3], g̊ʰ]. The "unvoiced voiced" symbols are used because lenis pronunciation of the unaspirated stops makes them similar to voiced stops in other language (Grønnum 1998:16).

3.1.2 Danish alveolar obstruents

Danish has three alveolar obstruent phonemes /d̥ tˢ s/. /tˢ/ is of primary focus in this paper. As opposed to the other aspirated plosives, the majority of its release phonetically consists not of aspiration, but of affrication. It is not an affricate, as that analysis would leave an unexpected gap in the phoneme inventory, and because that is hardly in accordance with the cognitive reality of the native speakers. It is phonologically an aspirated consonant in which the majority of the aspiration occurs with the tip of the tongue in close proximity to the alveolar ridge, acoustically resulting in affrication. The affrication is sometimes followed by a short period of regular aspiration after the tongue leaves the alveolar ridge, depending on the quality of the following vowel. The change from affrication to aspiration is gradual. Nothing in the recordings made for this paper or the existing literature suggests that this affrication is a dialectal, regiolectal, or gender-specific feature. Likewise, nothing in the recordings or the existing literature suggests that affrication of /tˢ/ in simple onset depends on the quality of the following vowel.

In the recordings made for this paper[4], the mean Voice Onset Time (VOT) of Danish /tˢ/ in simple onset (henceforth DA-*t*) lasted 94ms. For the average token, this consisted of 72ms of

[3] Depending on the quality of the following vowel
[4] See Sections 4 and 5 for recording and analysis methodology

affrication followed by 21ms of aspiration. Approximately half of the tokens (50%, n=138) consisted of no or only very short aspiration. In contrast, only 1% (n=3) of tokens had no affrication, and these tokens were all uttered by the same speaker. These numbers are similar to the results of Mortensen & Tøndering (2013:52), who find an average VOT of between 83.8ms and 86.8ms depending on subsequent vowel height. The relatively shorter VOT found in that paper is assumed to be due to their corpus consisting of spontaneous speech.

Table 2: DA-t VOT

Mean VOT	Mean fr.	Mean asp.	% without fr.	% without asp.	Range (cs)
93ms	72ms	21ms	1% (n=3)	50% (n=138)	3 – 17

3.2 SC

3.2.1 SC consonants in onset position

Following San (2007:24) the initial consonant phonemes of SC can be seen in Table 3. When Pinyin transcriptions differ from the phonetic symbols, they will be given in orthographic brackets <> following the phoneme. San does not consider the palatals to be phonemic.

Table 3: SC consonant phonemes in onset position

	Labial	Dental	Retroflex	(Palatal)	Velar
Unaspirated plosive	p <b>	t <d>			k <g>
Aspirated plosive	pʰ <p>	tʰ <t>			kʰ <k>
Unaspirated affricate		ts <z>	tʂ <zh>	tɕ <j>	
Aspirated affricate		tsʰ <c>	tʂʰ <ch>	tɕʰ <q>	
Fricative	f	s	ʂ ʐ <sh r>	ɕ <x>	x <h>
Nasal	m	n			
Lateral		l			

San's phoneme inventory of SC is minimal. He considers the glides /w j ɥ/ variants of the high vowels. According to Chao (1968:21), the velar fricative is a uvular /χ/. Chao describes /ʐ̩/ as a voiced continuant, but transcribes it phonemically using a rhotic /r/.

Similar to Danish, SC stops are not distinguished by voicing, but by aspiration. Unlike Danish, the unaspirated SC stops are lenis and the aspirated ones are fortis. The lenis series /p t̬ k ʈ̬ ʈʂ̬ tɕ/ could reasonably have been transcribed following the Danish "unvoiced voiced stop" convention /b̥ d̥ g̊ d̥s d̥ʒ̬ d̥ʑ/, but I am aware of no such convention in Chinese linguistics. The phonological status of the palatals [tɕ tɕʰ ɕ] is unsettled; while Cheng (1973:40) gives them phonological status, they have also been identified as allophones of the dental affricates (Hartman 1944:38) and the velar stops (Chao 1968:21), and San (2007:31) analyzes them as consonant-glide combinations, specifically combinations of the dental sibilants and palatal glide. Kratochvil (1968:25) describes the dental stops as alveolars, but his claim appears to be unfounded, while a dental place of articulation is backed up by x-ray and palatographic data (Zhou & Wu 1963, cited in San 2007:25).

3.2.2 SC dental obstruents

The main focus of this paper is the dental obstruent series /t̪ t̪ʰ ts̪ ts̪ʰ s/. Henceforth, italicized Pinyin symbols for the aspirated dental stop and dental affricates will be used instead of the phonetic symbols. The Pinyin symbols denote SC phonemes, not graphemes or phonemes of any other language.

In the recordings of native SC speakers made for this paper, the mean VOT of *t* lasted 86ms, consisting either of pure aspiration or velar frication before back vowels. This is slightly shorter than Danish, but longer than English (Lisker & Abramson 1964:394), even though all three are grouped together in Cho & Ladefoged's VOT typology (1999:223). The VOT mean is similar to that found by Chao & Chen (2008:223), who measured a mean VOT of 81ms for *t*, but markedly shorter than that found by Rochet and Fei (1991:105), who measured a mean VOT of 98.7ms for *t*. This difference is presumably because Rochet and Fei measured single syllables. The mean release duration of *z* in the recordings for this paper lasted 73ms, consisting solely of frication. These numbers are similar to Liu and Jongman (2013:4) who report a release between 60ms and 70ms[5]. The mean release duration of *c* in the recordings for this paper was 152ms, consisting of roughly two thirds affrication and one third aspiration (99ms and 53ms on average, respectively). Similar to *t*, the aspiration was often realized as velar frication before back vowels. 11% of the tokens had no audible aspiration; these were the tokens occurring in syllables with a syllabic voiced continuant /z̩/ in the rhyme, which is to be expected, as the place of articulation of this continuant is homorganic to that of the affricate (e.g Hartman 1944:31). The mean release duration of *c* is similar to that found by Liu and Jongman (2013:4). The release duration values in the recordings made for this paper are considered similar to previous findings.

Table 4: SC phonemes release duration

Phoneme	Mean VOT	Mean fr.	Mean asp.	% without fr.	% without asp.	Range (cs)
t	86ms	0ms	86ms	100% (n=112)	0% (n=0)	3 – 15
z	73ms	73ms	0ms	0% (n=0)	100% (n=98)	2 – 15
c	152ms	99ms	53ms	0% (n=0)	11% (n=8)	8 – 29

[5] Exact numbers are not given.

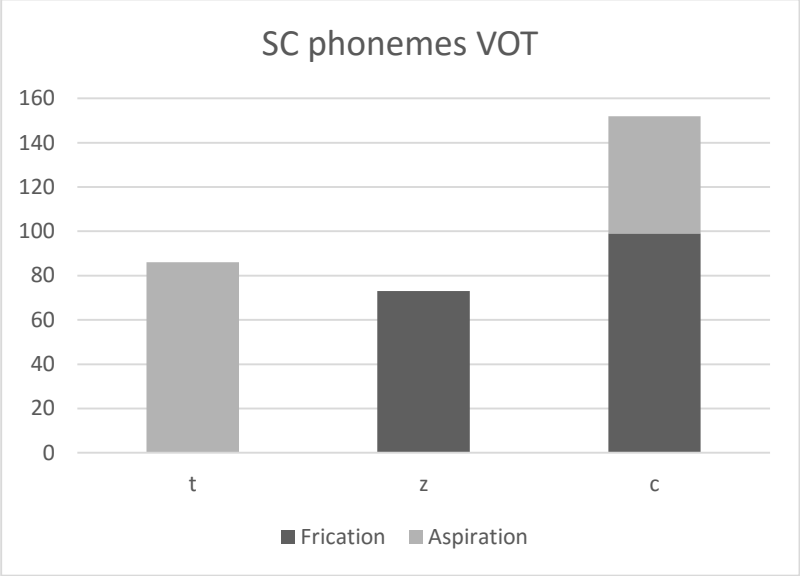

Figure 1: Illustration of the duration and contents of the releases of SC dental aspirated stop and affricates

3.3 Mapping expectations

There are plenty of minor differences in the Danish and SC simple onset consonants. Following the PAM and the SLM, many of these are likely to lead to different degrees of accented pronunciation of SC by native speakers of Danish.

/p k m f/ are expected to map to the corresponding sounds in Danish, which should be unproblematic, as the sounds are very similar. /pʰ kʰ/ are also expected to map to the corresponding sounds in Danish, the only difference being that the L1 sounds have lenis pronunciation. Given that the respective VOTs are similar (Mortensen & Tøndering 2013; Chao & Chen 2008), fortis pronunciation is not expected to be picked up during SC acquisition. /t s n l/ are alveolar in Danish and dental in SC, which should also be unproblematic, as the acoustic differences are minimal. Regarding the stop /t/, Cho and Ladefoged (1999:220) find that dental and alveolar places of articulation are not in themselves expected to result in differences in VOT, but Danish /d̥/ has been found to be up to 10ms longer than Chinese /t/ (Mortensen & Tøndering 2013:51; Chao & Chen 2008:223).

The nearest Danish equivalent to both the retroflex and palatal sets /tʂ tʂʰ ʂ tɕ tɕʰ ɕ/ is the alveolopalatal set [dʑ tɕ ɕ], phonologically considered to be /tˢj d̥j sj/ by Grønnum (1998:264-266). In PAM terms the difference here is assumed to be one of Category Goodness, with the SC palatals being the better match. Lai (2009:1268) found that perceptive discrimination of retroflex and palatal affricates were generally unproblematic for native speakers of Malay and Burmese, both of which only have postalveolar affricates. The SC rhotic /ʐ/ is expected to be sufficiently dissimilar to any Danish phoneme to form a novel category, unless orthographic interference triggers the Danish rhotic /ʁ/.

The SC velar fricative /x/ does not fit neatly into any Danish phoneme category. Pinyin interference may trigger /h/ for some Danish learners, particularly since /h/ is probably the nearest Danish category, although a very poor fit. Generally /x/ is expected to be sufficiently dissimilar

to form a novel category. Alternatively, many may already have a category /x/ due to experience with German.

Each of *t z c* are fairly close to DA-*t* in different respects, making it either a case of Single Category assimilation or a Category Goodness difference in PAM terms. In the following paragraphs, each will be categorized with regards to their similarities and dissimilarities to DA-*t*, setting up hypotheses for how native language transfer may affect pronunciation of them.

t is the closest to DA-*t* in duration. The complete lack of affrication makes it initially seem very different from DA-*t*, but since the affrication is hardly of cognitive significance to native speaker of Danish, it may not be of much importance to the transfer. Danish native speakers are also assumed to maintain the feature of affrication when pronouncing alveolar plosives /tʰ/ in other L2s (see Section 6). Due to orthographic interference, there is also a chance that written Pinyin *t* triggers [tˢ] even if *z c* are phonetically closer to DA-*t*.

Inspecting the VOTs uncritically, *z* appears very close to DA-*t*. The mean VOTs are reasonably similar, and considering that 50% of DA-*t* tokens have no audible pure aspiration, those 50% of the tokens should be very similar to the native *z* pronunciation. However, they are not so similar after all. *z* never has aspiration, and the affrication of *z* and that of DA-*t* are not identical, due to the gradual shift of DA-*t*'s affrication to more aspiration-like noise; the affrication noise of DA-*t* peaks right after the plosive release, while the affrication noise of the *z* release is more constant. The affrication noise of *z* also generally peaks at a higher frequency than that for DA-*t*. An unrelated reason for Danish speakers to be more likely to recognize *z* as a separate category is that German has a similar alveolar affricate /ts/ in syllable-initial position (Antonsen 2007:26). Finally, the relatively weak plosive release and fairly short duration of *z* means that it may be perceived as closer to Danish /s/ than DA-*t*. In this case both *z s* would content for the Danish /s/ category, with *z* being the weaker fit. The results point toward the latter analysis.

c is much longer than DA-*t*, but the realization is much more similar than the other contenders, as it is the only one of the consonants in which both affrication and aspiration are of importance. Aspiration is more prominent in *c*, and unlike DA-*t*, the shift from affrication to aspiration in *c* is not gradual. But the distribution of affrication and aspiration makes *c* phonetically very similar to Danish DA-*t*.

t c are about equidistant to DA-*t*, being similar to it in very different ways. An outcome of this could be that the early Chinese interlanguage of Danish L1 speakers will have *t c* grouped in a single L2 category, /tˢ/, which will have to be split at a later stage. Otherwise, both *t c* are expected to retain features from DA-*t*, i.e affrication is assumed to be retained in *t* and a gradual shift from affrication to aspiration is assumed to be retained in *c*, which would make them phonetically very similar. *z* is not assumed to be an equally good candidate, but could turn out to be grouped with the other two in the early interlanguage, especially considering Ne et al.'s (forthc.) finding that discrimination of *c z* by Danish listeners was generally harder than either *t c* or *t z*. These predictions were tested through recordings of a reading task done by Danish students of Chinese.

4 Material

4.1 Speakers

25 undergraduate students participated in the experiment. All students were native speakers of Danish and enrolled in the Chinese Studies program at Aarhus University. All except two participants were between 20 and 25 years old. The two others were 31 and 42 years old, respectively. More than half of the participants were from Central Jutland, the remainder coming from all across Denmark. 11 were male and 14 were female, which reasonably reflects the gender distribution of the program. All of the participants spoke English, all except two had studied German for at least three years in elementary school, and approximately half had studied French in high school. 10 of the participants had studied Chinese in high school, 9 more reported having some knowledge of Chinese before-hand, and 6 reported having no previous knowledge of Chinese whatsoever. The subjects were not compensated financially and did not receive credit for participating. None of the participants reported hearing, vision, or reading deficiencies. They were unaware of the particular hypothesis being tested. The subjects were at three different levels of their studies:

1st year students (henceforth Y1). 7 of the participants had begun their studies approximately 5 weeks previously. At this point they were reasonably proficient at using Pinyin, but only had a small inventory of Chinese characters and a small vocabulary.
2nd year students (henceforth Y2). 12 of the participants had begun their studies a little more than one year previously. At this point they had a reasonably large inventory of Chinese characters and vocabulary. They had been introduced to all characters and words in the experiment.
3rd year students (henceforth Y3). 6 of the participants had begun their studies a little more than two years previously. As part of their studies they had recently spent a semester in Beijing, where teaching had been done solely in SC, so they had been exposed to a large amount of both spoken and written Chinese.

In addition to the Danish participants, 7 native speakers of SC were recorded. 4 were male and 3 were female. They were all studying in Aarhus as exchange students. They were from different of areas of China, but all spoke SC as their native variety of Chinese.

No other factors than study level had significant effects.

4.2 Stimuli

The stimuli consisted of 64 simple sentences of 6 syllables each. 12 of the sentences were in Danish in order to get samples of the participants' L1 speech; 12 were in English; 40 were in SC. The Chinese participants were only asked to read the Chinese sentences. The Danish and English sentences were both designed to include the aspirated alveolar stop in front of a broad variety of vowel combinations. For both of these sets, 9 sentences included the alveolar stop and 3 were filler sentences. The Chinese sentences were designed to include the target dental obstruents before a broad variety of vowel and tone combinations. 30 of the sentences included one or more of these consonants while 10 were filler sentences, giving 47 critical stimuli.

The sentences used for the experiment can be found in Appendix 1.

4.3 Procedure

The participants were asked to read the sentences in Appendix 1 from slides of a Microsoft PowerPoint presentation. All sentences were presented on single slides, the Chinese ones in both characters and Pinyin. The recordings were self-paced and lasted 3-6 minutes on average for all sentences. The speech was recorded with a Zoom APQ3HD and a Sony PCM-D50. The recordings mostly took place in the Interacting Minds Centre lab at Aarhus University, but also in three other empty, quiet meeting rooms at Aarhus University and VIA University College respectively.

5 Analysis methodology

To investigate how L1 Danish speakers' pronunciation of dental obstruents differed from native pronunciation in SC, the software Praat (Boersma & Weenink 2005) was used to measure and analyze the release portion of the stops and affricates. The term VOT will be used to broadly signify the period of time from the plosive release until the beginning of voicing, thus covering the entire release of plosives and affricates. For each of the target dental obstruents, the VOT was measured and, if applicable, split into an affrication segment and an aspiration segment, which were measured independently. If either affrication or aspiration lasted less than 10ms they were considered inaudible, and thus given as 0ms. If consonants were mispronounced in a manner that suggested they were speech errors, they were excluded from the analysis. This mainly covers consonants pronounced at the wrong place of articulation, e.g dentals that were pronounced as retroflexes or velars. Dentals pronounced as alveolars were not excluded.

Individual segments were measured in centiseconds. While cs measurements are not particularly precise, VOTs proved hard to analyze using more specific measurements. For Chinese speakers, delimitation of affrication and aspiration was generally clear and unproblematic, and could reasonably have been given in ms. But for Danish speakers, more precision proved problematic, since the transition from affrication to aspiration in Danish pronunciation is much more gradual. This gradual transition was prominent among Danish L1 participants no matter which language they spoke. Delimitation of affrication and aspiration was based on spectrographic and auditory inspection of the individual tokens; if possible, high-pitched irregular noise was generally taken to mean affrication while pre-voicing formant traces was taken to mean aspiration. In the Danish speaker tokens, these frequently overlapped, in which case delimitation was based mainly on auditory judgment. An example of this can be seen on Figure 2, where unvoiced vowel formants can clearly be seen simultaneously with diminishing high-pitched irregular noise. When voicing did not start simultaneously for all formants, voice onset was taken to be at the beginning of voicing at F1.

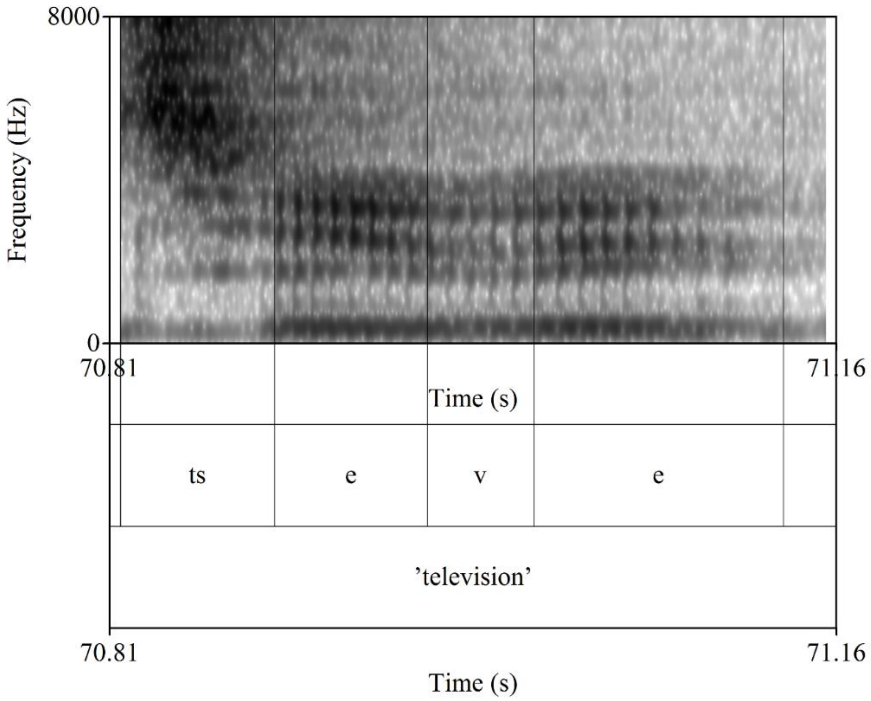

Figure 2: Unvoiced vowel formants simultaneous with diminishing fricative noise

6 Affrication in English alveolar stops

As mentioned above, it is often heard that Danes tend to pronounce /tʰ/, fortis /t/ and similar phonemes in different L2s as [tˢ], even in languages they are very proficient in. If true, this is assumed to increase the likelihood for Chinese *t* to be pronounced as [tˢ]. To test it, a subset of the current experiment tested the pronunciation of the English aspirated alveolar stop by L1 speakers of Danish. The participants all spoke English with a high proficiency. 25 students of Chinese cannot be said to be representative for the whole population of English-speaking Danes, so work is still needed on this topic, but the results are sufficiently clear to have some merit. No native English speakers were recorded for comparison, so comparison will be made to the VOTs reported in Docherty (1992:116) and Klatt (1975:689). They find mean VOTs for /tʰ/ at 63ms and 65ms for British English and American English respectively. None of them refer to affrication in these stops, so the VOT reported in those papers is given as aspiration only in Table 5 and Figure 3.

It turned out there was a small, but significant, difference in VOT of DA-*t* and English /tʰ/ as pronounced by Danish speakers (henceforth EN-*t*), with $t(523)=2.6$, $p<0.01$. The mean EN-*t* had a VOT of 88ms, which is just 6ms shorter than DA-*t*, and still around 20ms longer than native pronunciation. Likewise, affrication was slightly less prominent in EN-*t* than in DA-*t*; EN-*t* had a mean affrication duration of 66ms, thus an average of 74% of the release (comp. 77% for DA-*t*). In total, affrication was avoided in 6% (n=16) of EN-*t* tokens, which is a lot more frequent than DA-*t*, but still means that there is affrication in the vast majority of tokens. Half of the tokens with no affrications (n=8) were spoken by two students. 48% (n=120) of EN-*t* tokens has no period of aspiration proper, which is more or less the same as for DA-*t*.

*Table 5: EN-*t *vs DA-*t *VOT*

Phoneme	Mean VOT	Mean fr.	Mean asp.	% without fr.	% without asp.	Range (cs)
EN-*t*	88ms	66ms	23ms	6% (n=16)	48%(n=120)	4 - 16
DA-*t*	93ms	72ms	21ms	1% (n=3)	50% (n=138)	3 - 17

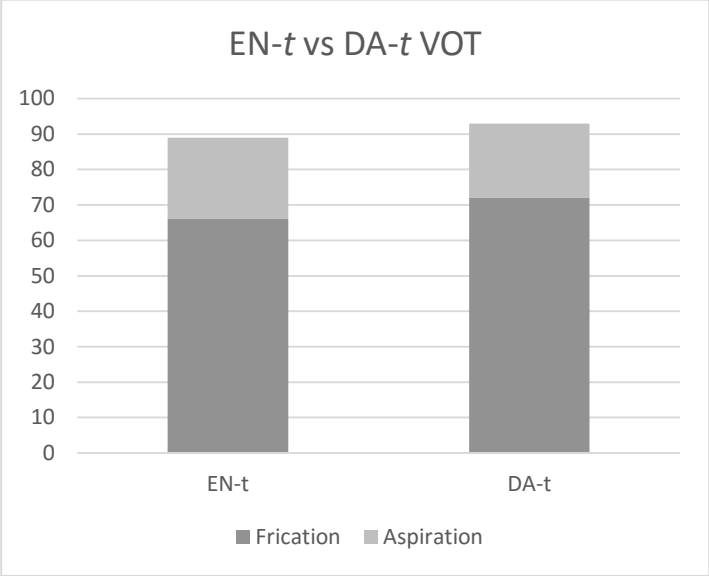

*Figure 3: Illustration of the duration and contents of the EN-*t *and DA-*t *release*

In conclusion, EN-*t* is remarkably close to DA-*t* in pronunciation, and it can safely be said that affrication has been retained as a prominent feature in EN-*t*. As with DA-*t*, the feature is prominent for *all* recorded speakers – including those who consistently avoided it in their pronunciation of Chinese *t* (see below). This means that even those students who had successfully established /tʰ/ as a stable category in SC retained [tˢ] in their pronunciation of English.

7 Results

7.1 t

Unlike *z c*, *t* can appear before /i/ and can be palatalized, and both /t̪ʰi/ and initial /t̪ʰj/ were involved in the experiment. They were, however, left out of the final analysis for two reasons: 1) since the dental affricates cannot appear before high front vowels, /tʰ/ will most likely not be perceived as a non-affricate whether it has affrication or not, and 2) aspiration preceeding a high front vowel is acoustically very similar to affrication, making delimitation of either very difficult and ultimately too random. The latter proved true for both Danish and native Chinese speakers.

t was the only analyzed consonant in which the subjects' total VOT became progressively *less* native-like as they advanced in their studies. The mean total VOT for Y1 students was at 92ms, for Y2 students 95ms, and while both were longer than the mean VOT for native Chinese

speakers, neither differed significantly from it. The Y3 students, however, had a mean total VOT at 112ms, which was significantly longer than the native speakers, with $t(205)=5.34$, $p<0.001$. It should be mentioned that the Y1 and Y2 students' VOT were not significantly different than the VOT of DA-*t* either, while the Y3 students' were, with $t(368)=5.49$, $p<0.001$. It should also be mentioned that the long VOT for Y3 students can be partially explained by one of the participants of this group having a particularly long mean VOT of 152ms, raising the Y3 mean a fair bit. However, if the outlier participant is excluded, the Y3 mean VOT is at 104ms, which remains significantly higher than both the Y1, Y2, and native Chinese groups.

The average Y1 *t* consists of 41ms of affrication followed by 50ms of aspiration. None of the tokens actually look like this; this average is a product of a very variant pronunciation. More than half (65%, n=73) of the tokens had affrication, and among those the mean affrication duration was 64ms. Out of the Y1 participants, 4 affricated all or almost all of their tokens, 1 affricated approx. half of his/her tokens, and 2 almost never affricated. A third (30%, n=34) of the tokens had less than 10ms or no period of aspiration. The majority of Y1 *t* VOTs consisted of a fairly long period of affrication sometimes followed by a fairly short period of aspiration.

The average Y2 *t* consists of 24ms of affrication and 71ms of aspiration. This is also a product of variant pronunciation, but clearly shows that by this study level affrication has become much less prominent in the pronunciation of *t*. Less than half (43%, n=83) of Y2 tokens had affrication, while only 9% (n=18) lacked aspiration. None of the Y2 students affricated all of their tokens, but only 2 out of 12 consistently avoided it. The majority of the Y2 *t* VOTs consisted either only of aspiration or of a fairly short period of affrication followed by a longer period of aspiration.

The average Y3 *t* consists of 31ms of affrication and 81ms of aspiration. Similar to Y2, 44% (n=42) of tokens had affrication, and 11% (n=10) lacked aspiration. These numbers are less native-like than the Y2 ones. There are two explanations to account for this development: 1) a longer VOT makes affrication a less prominent part of the release and 2) a longer VOT is a product of the students developing fortis pronunciation of the phoneme. A more fortis pronunciation of a Danish spirantized *t* would acoustically result in a longer release retaining both affrication and aspiration. An adverse effect of this strategy is that the prolonged VOT causes these tokens to resemble *c* more than *t*. For an extreme example of this, see Figure 4, which is a *t* token with a very long release (29cs) that has both affrication and aspiration. Presumably the prolonged VOT is an attempt to stress the aspiration of the release, but in doing so, affrication also becomes more prominent, and the token altogether comes to resemble native *c* more than *t*.

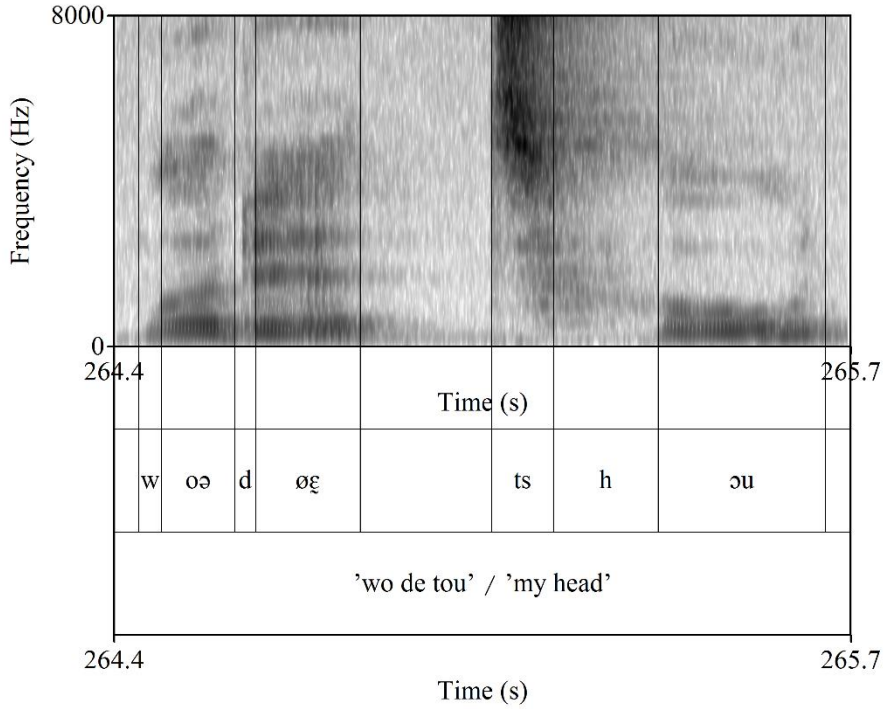

Figure 4: Example of very lengthy VOT of t *token with affrication and aspiration both being prominent*

Half of the Y3 students consistently or near-consistently avoided affrication, which is a clear improvement from Y2. The majority of the Y3 *t* VOTs had no affrication, while some consisted of both affrication and aspiration, but with fairly long VOTs, which both stresses the aspiration of these tokens, and results in a less native-like release duration.

Table 6: t *VOT*

Group	Mean VOT	Mean fr.	Mean asp.	% without fr.	% without asp.	Range (cs)
Y1	92ms	41ms	50ms	35% (n=39)	30% (n=34)	4 - 21
Y2	95ms	24ms	71ms	57% (n=108)	9% (n=18)	3 - 19
Y3	112ms	31ms	81ms	56% (n=53)	11% (n=10)	5 - 29
Native	86ms	0ms	86ms	100% (n=112)	0% (n=0)	3 - 15

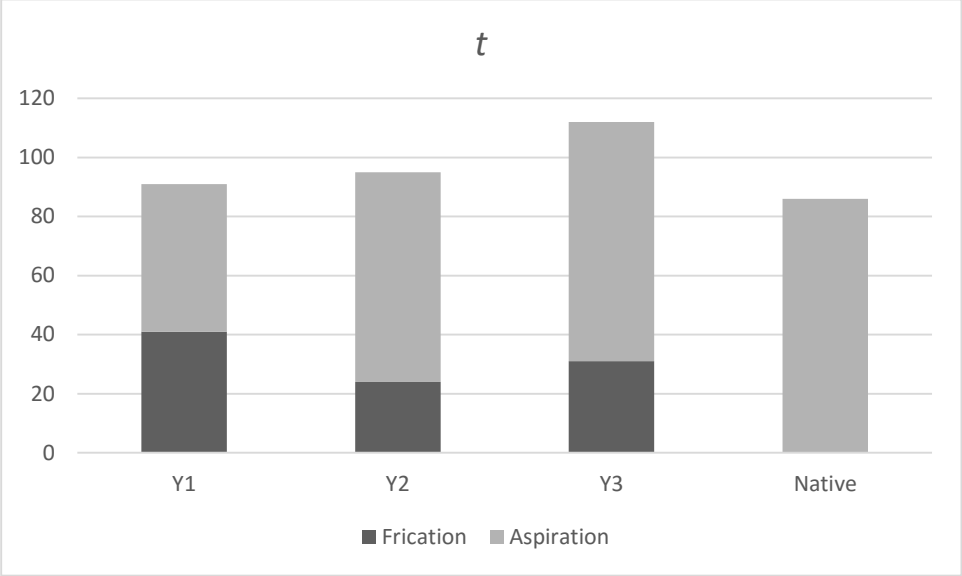

Figure 5: Illustration of the duration and contents of the t *release at different study levels and for native speakers*

7.2 z

The pronunciation of *z* became progressively more native-like as the students advanced in their studies. The mean total VOT for Y1 students was at 131ms. The Y2 students' VOT was significantly shorter at 119ms with $t(239)=2.64$, $p<0.01$. Likewise, the Y3 students' VOT was significantly shorter than the Y2 students at 106ms with $t(231)=2.06$, $p<0.05$. However, the Y3 students' VOT was still significantly longer than that of the native Chinese speakers, with $t(179)=6.10$, $p<0.001$. Except for a few mispronounced tokens, the participants never pronounced it with aspiration, and it is very clearly pronounced distinctly from *t* and *c*. The only difference between the study levels is the duration of the affrication. It is notable that while the duration gradually improves, the Y3 mean duration is still much longer than the native mean.

z often has no plosive release at all, making it phonetically very similar or identical to /s/. This was a very frequent issue, happening in 25% (n=84) of all Danish tokens. Some speakers never make this error, but 14 of the participants made it at least once. There was no significant improvement as studies progressed: between 20% and 30% of all tokens for a given group lacked plosive release. For an example, see Figure 6 showing a *z* token with no plosive release as said by a Y3 speaker. At 11cs, the token is near the average duration for Y3 students.

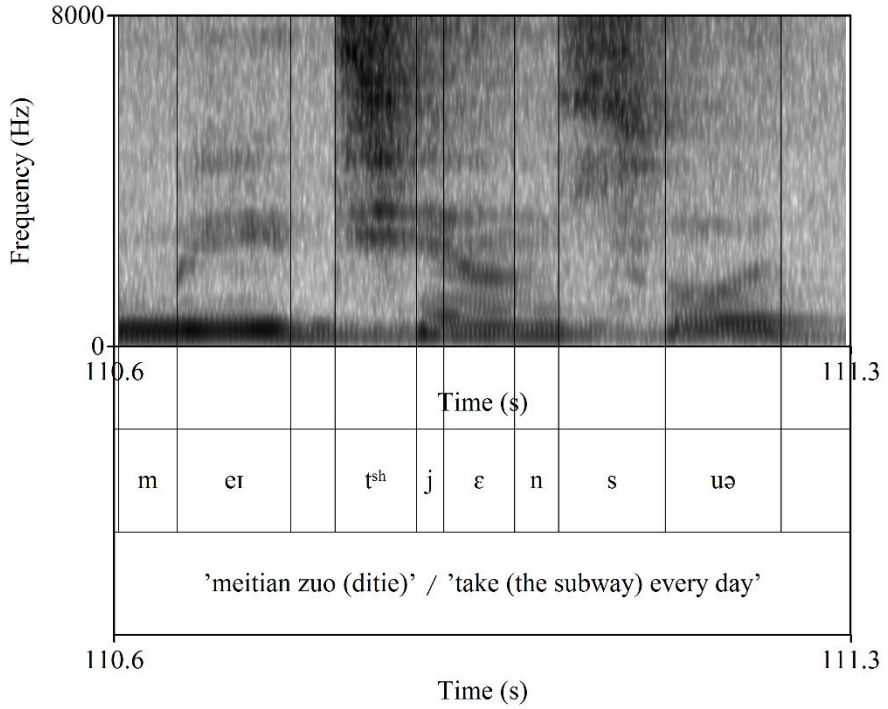

Figure 6: Example of a z *token with no plosive release*

It should be noted that a few of the native Chinese tokens also lacked plosive release, but these cases can all be accounted for by rapid speech, which cannot be said for the Danish speakers.

Table 7: z *VOT*

Group	Mean VOT	Mean fr.	Mean asp.	% w.o. fr.	% w.o. asp.	% w.o pl.rel.	Range (cs)
Y1	131ms	131ms	1ms	0% (n=0)	99% (n=90)	23% (n=21)	5 - 23
Y2	119ms	118ms	1ms	1% (n=1)	98% (n=160)	27% (n=45)	5 - 29
Y3	106ms	105ms	1ms	0% (n=0)	99% (n=82)	22% (n=18)	4 - 23
Native	73ms	73ms	0ms	0% (n=0)	100% (n=98)	4% (n=4)	2 - 15

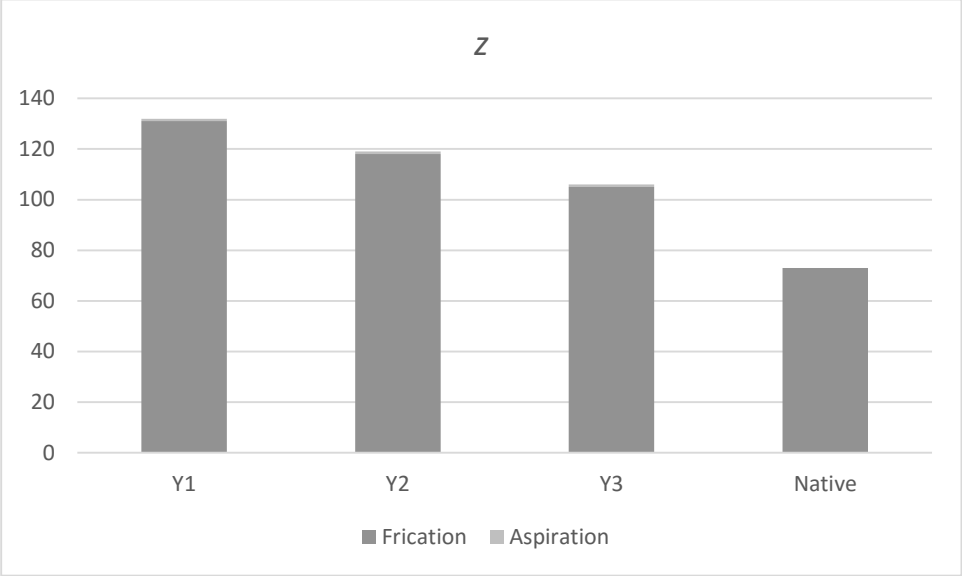

Figure 7: Illustration of the duration and contents of the z *release at different study levels and for native speakers*

7.3 c

The pronunciation of *c* became progressively more native-like as the students advanced in their studies. The VOT is not significantly different between Y1 and Y2 at 110ms and 109ms respectively. The Y3 mean VOT is significantly longer at 125ms, with $t(171)=3.13$, $p<0.01$. While still much shorter than native pronunciation, it is a clear improvement.

The average Y1 *c* consists of 72ms of affrication followed by 38ms of aspiration. The mean distribution is actually similar to the native Chinese pronunciation, with 65% of the VOT consisting of affrication, but again the average is a result of very variant pronunciation. This early in their studies, the participants clearly had troubles with the pronunciation of *c*; 14% (n=10) of the tokens were mispronounced to the point of exclusion. 45% (n=27) were pronounced with no or minimal aspiration, though a clear fortis pronunciation typically distinguished these from *z*. To further complicate the picture, 13% (n=8) were pronounced with no affrication, making these tokens more like native *t*, while still typically longer. This is presumably a form of hypercorrection, demonstrating the students' problems with keeping *t c* separate. An example of this can be seen on Figure 8, where a Y1 student pronounces a *c* token with no signs of affrication, but a fairly long period of aspiration (13cs).

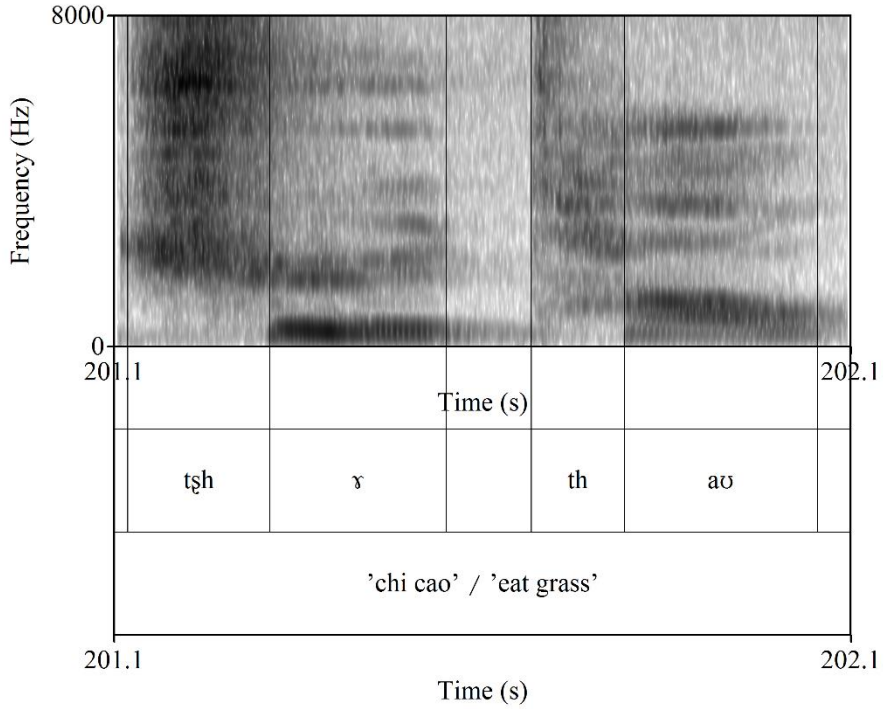

Figure 8: Example of lengthy c *token with no affrication*

The average Y2 *c* consists of 64ms of affrication followed by 45ms of aspiration. This means that the mean distribution is less native-like than at Y1, with only 59% of the average VOT consisting of affrication. This is actually a result of the pronunciation becoming more stable, and aspiration becoming more prominent: only 27% (n=31) of tokens were pronounced without aspiration. Along with much fewer tokens discarded due to mispronunciation (6%, n=7), it is clear that *c* pronunciation has stabilized and become less problematic. Lack of affrication was still surprisingly prominent, occurring in 17% (n=19) of Y2 *c* tokens.

The average Y3 *c* consists of 86ms of affrication followed by 39ms of aspiration. Along with duration, the distribution has also grown more native-like, with affrication lasting an average of 69% of the total duration. Lack of aspiration is still a significant issue, seen in a third (33%, n=20) of tokens. Lack of affrication has, however, ceased to be a significant issue (3%, n=2).

Table 8: c *VOT*

Group	Mean VOT	Mean fr.	Mean asp.	% without fr.	% without asp.	Range (cs)
Y1	110ms	72ms	38ms	13% (n=8)	45% (n=27)	3 - 25
Y2	109ms	64ms	45ms	17% (n=19)	27% (n=31)	5 - 20
Y3	125ms	86ms	39ms	3% (n=2)	33% (n=20)	6 - 23
Native	152ms	99ms	53ms	0% (n=0)	11% (n=8)	8 - 29

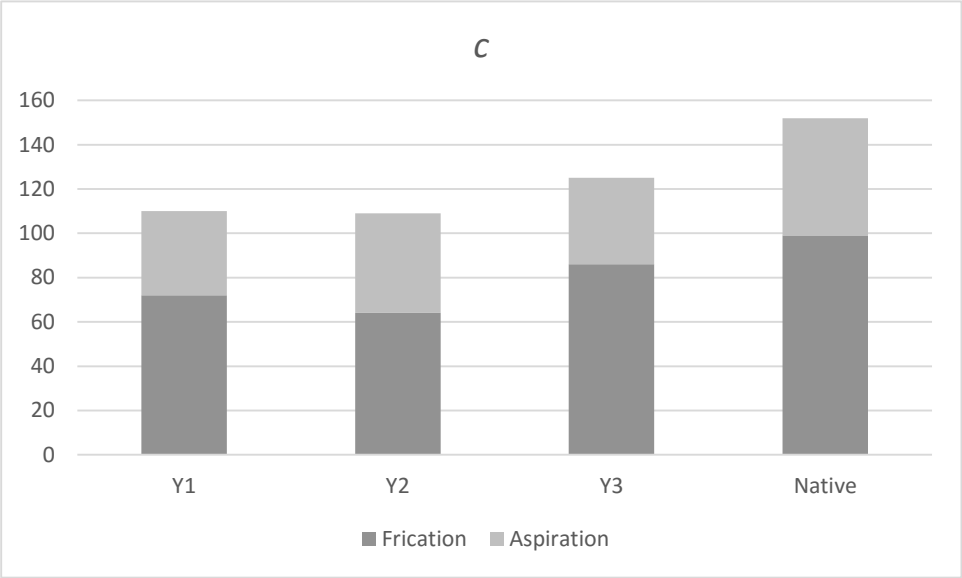

Figure 9: Illustration of the duration and contents of the c *release at different study levels and for native speakers*

8 Discussion

The results of the experiment presented in Section 6 convincingly showed that at least some participants had developed a novel category EN-*t* separate from DA-*t*, and that for all of these participants the difference between the two categories was one of duration and not one of affrication. It appears that affrication is a very stable feature of the DA-*t* category, which should pose a problem in SC acquisition. In some ways it does and in some ways it doesn't. Looking at the data in Section 7, it appears that different categories for *t z c* are actually established early on in SC acquisition.

The category for *z* in particular is established as distinct early on, being much longer than *t c* and never having aspiration. This may be because *z* is more likely to be perceived as /s/ rather than DA-*t*. In any case, plosive release is not as important a feature among the *z* of Danish speakers as it is among L1 speakers. Danish speakers may use other acoustic cues to distinguish the *z* from /s/; the recordings indicate that an initially very narrow constriction of *z* may be an important cue. The possibility that *z* might be categorically closer to /s/ than to DA-*t* was not taken into account when the experiment design was made, but a detailed comparison of *z* and /s/ as pronounced by Danish speakers would be very enlightening, and could also help provide a solid explanation for the very long release of *z* among Danish speakers. It would also be very interesting to see whether SC /s/ sometimes gets plosive release when pronounced by Danish speakers.

t c are also more or less established as separate categories for most or all speakers even by the 5th week of Chinese studies. This is corroborated by the fact that they were distinguished near-perfectly by the first year students tested in Ne et al (forthc.), and by the fact that *c* is 18ms longer than *t* on average for the Y1 group. But both categories show very differing pronunciation within-group, and there is a significant overlap of *t c* pronunciation. The *t* category appears

to still be identical to DA-*t* for some Y1 speakers. For one speaker *c* appears to be identical to DA-*t* as well. That is demonstrated in Figure 10, in which both *t c* appear in close proximity to each other with nearly identical pronunciation. They are distinguished only by what sounds like a slightly louder pronunciation of *c*, and a minor durational difference, though *c* is still well within range of *t* for the same speaker.

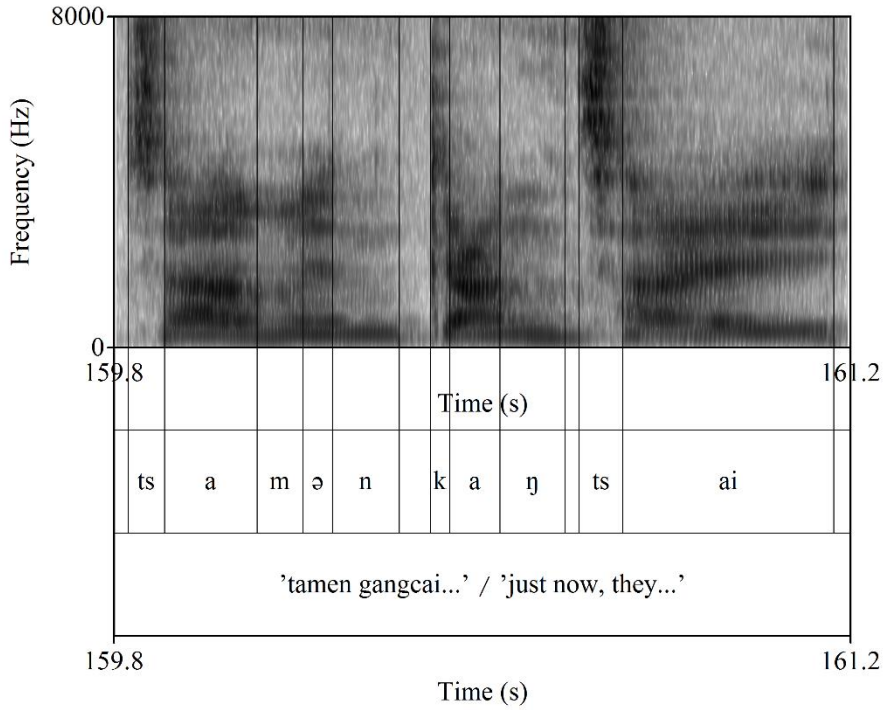

Figure 10: Example of two very similar t *and* c *tokens in close proximity to each other*

For many speakers at Y1 level, the distinguishing feature between *t c* appears to be one of intensity, and not affrication. Affrication is typically a stable feature of both of the categories, so they are distinguished instead by a more fortis pronunciation in *c*, which results in a longer release duration. Of course, in native SC, both of *t c* are fortis, but that may be hard for Danish listeners to perceive. Given the results by Ne et al. (forthc.) the categorization strategy works well for perceptual distinction, but for productive distinction it is problematic, considering the significant overlap it is likely to create between the two categories in casual speech.

At Y2 and Y3 level, *c* becomes progressively more native-like as its release duration increases, but it does not necessarily become progressively more distinct from *t*. Generally, *t* gradually develops a more fortis pronunciation, causing an increase in release duration, even though it was already long compared to native pronunciation. This is true for almost all participants in the experiment; at Y3, the individual participants have average release durations ranging from 97ms to 152ms for *t*, meaning that even the speaker with the shortest average VOT has a longer VOT than the average for all Y2 participants. For some speakers, affrication ceases to be a feature of *t*, and for some it doesn't. As far as productive distinction goes, this makes all the difference: if *t* is pronounced with a prolonged release duration but with no affrication, the native speaker of SC is expected to readily perceive it as *t*. But if it is pronounced with both a prolonged release duration *and* affrication, what is left to distinguish it from *c*? The release of

c is still longer on average, but only with 13ms, so a very significant overlap between the two is to be expected. As mentioned above, only half of the Y3 participants consistently avoided affrication, which means that this is a very tenacious issue.

While different categories for *t c* were established early on, they were relatively unstable for at least Y1 and Y2. A significant number of *c* tokens were realized simply as aspirated stops /tʰ/. There are two possible explanations for this: 1) It is due to hypercorrection among students who are particularly aware that they have to aspirate and not affricate their *t*s; the fact that this transfers to *c* shows that distinction between the two is still not entirely straightforward. 2) Affrication is not a distinguishing feature of either *t c* for the students who made the error; the non-affricate tokens typically have a long release, indicating duration or fortis pronunciation, which are important features for *c*. It may simply be that some students distinguish *t c* by duration or fortis pronunciation, and affrication is an optional feature of both.

As Flege (1981:446) points out, it is not at all clear whether or not "native-like pronunciation at the level of phonetic implementation is even necessary for accent-free speech". But in this case, the particular difference between phonetic implementation of the Danish /tˢ/ category and the SC /tʰ tsʰ/ categories may cause significant communicative problems. Further research is needed on the extent of these communicative problems, and the extent to which these sounds are perceived as the right or wrong categories by native speakers of SC remains unknown.

A solution to the problem may lie in the early teaching of SC phonetics. As mentioned above, affrication of DA-*t* may not be a very cognitively salient feature for the native speaker, and Danes are likely to be unaware of the feature. The feature will certainly be easier to eradicate from their Chinese pronunciation if they are aware of it. When teaching SC phonetics to Danish learners, teachers are advised to make students aware that affrication is a prominent feature of aspirated Danish stops and not of aspirated SC stops.

9 Conclusion

In spite of the Danish phonology being poorly equipped to help categorize the SC phonemes /tʰ ts tsʰ/ in a meaningful manner, Ne et al. (forthc.) showed that Danes had no significant problems with perceptual discrimination between them. This paper set out to test whether Danes could productively discriminate between the sounds.

The findings suggest that Danes generally make a productive distinction between the sounds even at a fairly early level of their studies, though the features used for the productive discrimination are problematic. SC /ts/ turns out to be productively more similar to Danish /s/ than /tˢ/, and while it is certainly a category in itself, Danes fairly often produce the affricate with no plosive release at the beginning, making it simply a fricative which sometimes has a very narrow constriction at the onset. Danish phonology makes native speakers of Danish likely to perceive and produce aspirated stops as lenis; an effect of this is that the main discriminating feature between SC /tʰ tsʰ/ for some Danes turn out not to be affrication, but duration – especially in the early Chinese interlanguage. Over time, fortis pronunciation of /tʰ/ is developed, which only results in a longer and less-native like release duration, especially since some speakers fail to develop the distinction between affrication and aspiration. At the third year of their studies, some participants consistently produced /tʰ/ with no affrication, while some still struggled with

pronouncing it distinctly from /tsʰ/. The findings are generally in line with the hypothesis made on the basis of the PAM and SLM models.

References

Antonsen, Elmer H. 2007. *Elements of German. Phonology and Morphology*. Tuscaloosa: The University of Alabama Press.

Bassetti, Benedetta. 2006a. Pinyin orthographic input and CFL teachers' pronunciation. Paper presented at *The British Chinese Language Teachers' Seminar, Cambridge, 9-10 September 2006*.

Bassetti, Benedetta. 2006b. Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy* 9(1). 95-114.

Bassetti, Benedetta. 2008. Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (eds.): *Input Matters in SLA*. Bristol: Multilingual Matters. 191-206.

Best, Catherine T.; McRoberts, Gerald W.; Sithole, Nomathemba M. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infnats. *Journal of Experimental Psychology: Human Perception and Performance* 14(3). 345-360.

Best, Catherine T.; McRoberts, Gerald W.; Goodell, Elizabeth. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109(2). 775-794.

Boersma, Paul; Weenink, David. 2005. *Praat: doing phonetics by computer* (Version 4.3.01.) [Computer program]. Retrieved from http://praat.org.

Caramazza, A.; Yeni-Komshian, G. H.; Zurif, E. B.; Carbone, E. 1973. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America* 54(2). 421-428.

Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley, Los Angeles & London: University of California Press.

Cheng, Chin-Chuan. 1973. *A synchronic phonology of Mandarin Chinese*. The Hague & Paris: Mouton de Gruyter.

Cho, Taehong; Ladefoged, Peter. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27. 207-229.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Berlin & New York: Foris Publications.

Flege, James Emil. 1980. Phonetic approximation in second language acquisition. *Language Learning* 30(1). 117-134.

Flege, James Emil. 1981. The phonological basis of foreign accent: a hypothesis. *Teachers of English to Speakers of Other Languages Quarterly* 15(4). 443-455.

Flege, James Emil. 1995. Second language speech learning. Theory, findings, and problems. In W. Strange (ed.): *Speech perception and linguistic experience: Issues in cross-language speech research*. Baltimore: York Press. 233-277.

Grønnum, Nina. 1998. *Fonetik og fonologi. Almen og dansk*. Copenhagen: Akademisk Forlag.

Grønnum, Nina. 2007. *Rødgrød med fløde. En lille bog om dansk fonetik*. Copenhagen: Akademisk Forlag.

Hartman, Lawton M. 1944. The segmental phonemes of the Peiping dialect. *Language* 20(1): 28-42.

Klatt, Dennis H. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18. 686-706.

Kratochvil, Paul. 1968. *The Chinese language today. Features of an emerging standard*. London: Hutchinson University Library.

Kuhl, Patricia A.; Williams, Karen A.; Lacerda, Francisco; Stevens, Kenneth N.; Lindblom, Björn. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255(5044). 606-608.

Lai, Yi-hsui. 2009. Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese. *Language and Cognitive Processes* 24(7/8). 1265-1285.

Li, Charles N; Thompson, Sandra A. 1981. *Mandarin Chinese. A functional reference grammar*. Berkeley & Los Angeles: University of California Press.

Lisker, Leigh; Abramson, Arthur S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20. 384-422.

Liu, Jiang; Jongman, Allard. 2013. American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/. *Proceedings of Meetings on Acoustics* 18.

Lively, Scott E.; Pisoni, David B. 1997. On prototypes and phonetic categories: a critical assessment of the Perceptual Magnet Effect in speech perception. *Journal of Experimental Psychology* 23(6). 1665-1679.

Meng, Z. 1998. 对外汉语语音教学中使用 "汉语拼音方案" 的几个问题 Duiwai hanyu yuyin jiaoxue zhong shiyong "hanyu pinyin fang'an" de jige wenti. In J. Chao & Z. Meng (eds.): 语音研究与对外汉语教学 *Yuyin yanjiu yu duiwai hanyu jiaoxue*. Beijing: Beijing Language and Culture University Press.

Mortensen, Johannes; Tøndering, John. 2013. The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish. In R. Eklund (ed.): *Proceedings of Fonetik 2013. The XXVIth Annual Phonetics Meeting, Linköping University, Linköping, Sweden.* Linköping: Linköping University. 49-52.

Ne, Xuechao; Yan, Mengzhu; Zhang, Chun; Sloos, Marjoleine. Forthcoming. The discrimination of <tai cai zai> by Danish listeners. In *Linguistics in Chinese Education*. Beijing: Beijing Language and Culture University Press.

Rochet, Bernard L.; Fei, Yanmei. 1991. Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception. *Canadian Acoustics* 19(4): 105-106.

San Duanmu. 2007. *The phonology of Standard Chinese*. 2nd ed. Oxford & New York: Oxford University Press.

Sloos, Marjoleine; Liang, Jie; Ne, Xuechao; Hansen, Rasmus Puggaard; Yan, Mengzhu; Zhang, Chun. 2015. Speech rhythm of the Danish-Chinese interlanguage relies on rhyme structure. In *Proceedings of the Oriental COCOSDA/CASLRE Conference, October 28-30 2015, Shanghai*. Shanghai: Jiao Tong University. 305-310.

Sloos, Marjoleine; Liang, Jie; Yan, Mengzhu; Zhang, Chun. Forthcoming a. Acquisition of pitch in Chinese by Danish learners. In *Proceedings of the 2nd Workshop "Chinese Accents and Accented Chinese", 26-27 October 2015, Nordic Centre, Fudan University, Shanghai*.

Sloos, Marjoleine; Yan, Mengzhu; Liang, Jie; Zhang, Chun. Forthcoming b. Danish acquisition of Mandarin Chinese tones.

Wang Lei. 2014. Danish accented Chinese: some predictions. Paper presented at *Chinese Accents and Accented Chinese Workshop*, Nordic Centre, Fudan University, Shanghai, 9-10 October 2014.

Yang, Bei. 2015. *Perception and production of Mandarin tones by native speakers and L2 learners*. Berlin & Heidelberg: Springer.

Zhou Dianfu; Wu Zongji. 1963. 普通话发音圖譜 *Putonghua fayin tu pu*. Beijing: Shangwu Yinshuguan.

Appendix 1: Sentences used in experiment

The sentences used in the experiment for this paper can be seen below. The Danish sentences are accompanied by glosses and translations. The Chinese sentences are written in both Pinyin and characters, and are accompanied by glosses and translations. Explanations for the abbreviations used in the glosses can be found in Appendix 2.

Danish sentences:

DA-1 Mi-t kinesisk er god-t
1sg.GEN-NEU Chinese be.PRS good-NEU
My Chinese is good

DA-2 Det bliv-er ti kron-er
3sg.NEU become-PRS 10 crown-PL
That will be ten Crowns

DA-3 Du skal tal-e med ham
2sg.NOM must.PRS speak-INF with 3sg.MASC.OBL
You have to speak with him

DA-4 Han se-r altid tv
3sg.MASC.NOM see-PRS always tv
He always watches tv

DA-5 Jeg spis-er jævnlig-t kiks
1sg.NOM eat-PRS regular-ADVZ crackers
I regularly eat crackers

DA-6 Hvor ligg-er det høj-e tårn?
where lie-PRS DEF.NEU high-DEF tower
Where is the high tower?

DA-7 Hun se-r altid Netflix
3sg.FEM.NOM see-PRS always Netflix
She always watches Netflix

DA-8 Den ide er altså træls
DEF.NEU idea be.PRS really annoying
Really, that idea is annoying

DA-9 Hvornår er det min tur?
when be.PRS 3sg.NEU 1sg.GEN.COM turn
When is it my turn?

DA-10	Aftens-mad-en		var	lækker
	evening-food-DEF.COM		be.PST	delicious
	The dinner was delicious			

DA-11	Husk	at	børst-e	tænd-er
	remember	INF	brush-INF	tooth-PL
	Remember to brush your teeth			

DA-12	Han	er	lidt	tvetydig
	3sg.MSC.NOM	be.PRS	a.little	ambiguous
	He's a little ambiguous			

English sentences:

EN-1	Learning Chinese is hard
EN-2	Two friends meet in a bar
EN-3	Don't talk about politics
EN-4	We all like jazz music
EN-5	My t-shirt is ruined
EN-6	Thailand is beautiful
EN-7	That'll be ten dollars
EN-8	I don't own a cell phone
EN-9	I enjoy sun tanning
EN-10	Will you tell me the truth
EN-11	The party is tonight
EN-12	Japan is beautiful

Standard Chinese sentences

SC-1	孩子常常踢球			
	háizi	chángcháng	tī	qiú
	child	often	kick	ball
	The child often kicks the ball			

SC-2	总之你们还好		
	zǒngzhī	nǐ-men	háihǎo
	anyway	2p-PL	not.bad
	Anyway, you're fine		

SC-3	汉字写得不错			
	hàn-zì	xiě	de	bù-cuò
	Han-character	write	COMPL	NEG-wrong
	You're not bad at writing Chinese characters			

SC-4 词典比较便宜
cídiǎn bǐjiào piányi
dictionary compare cheap
The dictionary is relatively cheap

SC5 我每天坐地铁
wǒ měi-tiān zuò dìtiě
1p every-day sit subway
I take the subway every day

SC6 这烤鸭很好吃
zhè kǎo-yā hěn hǎo-chī
this roast-duck very good-eat
This roast duck is very tasty

SC7 他坐出租汽车
tā zuò chūzū-qìchē
3p sit rent-car
He takes a cab

SC8 王朋喜欢跳舞
Wáng Péng xǐhuan tiàowǔ
Wang Peng(name) like dance
Wang Peng likes to dance

SC9 他们刚才起床
tā-men gāngcái qǐ-chuáng
1p-PL just rise-bed
They just got out of bed

SC10 太太吃我的醋
tàitai chī wǒ de cù
wife eat 1p SBORD vinegar
My wife is jealous of me

SC11 她在餐厅吃饭
tā zài cān-tīng chī fàn
3p at meal-hall eat food
She eats at a restaurant

SC12 咱们不是老师

zánmen bù shì lǎoshī
3pl.INCL NEG be teacher
We are not teachers

SC13 我弟弟是老师

wǒ dìdi shì lǎoshī
1p younger.brother be teacher
My younger brother is a teacher

SC14 身体健康很好

shēntǐ jiànkāng hěn hǎo
body health very good
My health is very good

SC15 你常常说 ”糟糕”

nǐ chángcháng shuō zāogāo
2p often say rotten-cake
You often say 'crap'

SC16 我有一把钥匙

wǒ yǒu yī bǎ yàoshi
1p have one bundle key
I have one key

SC17 我跟朋友讨论

wǒ gēn péngyou tǎolùn
1p with friend discuss
I discuss it with a friend

SC18 羊在山上吃草

yáng zài shān shàng chī cǎo
sheep at mountain on eat grass
The sheep eats grass on the mountain

SC19 老师找图书馆

lǎoshī zhǎo túshūguǎn
teacher look.for library
The teacher is looking for the library

SC20 我有两把椅子

wǒ yǒu liǎng bǎ yǐzi
1p have pair bundle chair
I have two chairs

SC21 英文很难学好

yīng-wén hěn nán xué-hǎo
English-language very difficult study-good
It is hard to learn speaking English well

SC22 你的工作太忙

nǐ de gōngzuò tài máng
2p SBORD work(N) too busy
You're too busy working

SC23 我们谈了一下

wǒ-men tán le yīxià
1p-PL talk PRF for.a.little.while
We talked for a little while

SC24 学不学都可以

xué bu xué dōu kěyǐ
study NEG study all okay
Studying and not studying are both okay

SC25 老师开始上课

lǎoshī kāishǐ shàng-kè
teacher begin up-lesson
The teacher begins the lesson

SC26 你喜欢吃白菜

nǐ xǐhuan chī bái-cài
2p like eat white-vegetable
You like eating Chinese cabbage

SC27 学生都很聪明

xuésheng dōu hěn cōngming
student all very clever
The students are all clever

SC28 不要托运行李

bú yào tuō-yùn xíngli
NEG want trust-carry luggage
I don't want to check in my luggage

SC29 中文你说一说

zhōng-wén nǐ shuō yi shuō
Chinese-language 2p talk one talk
Speak a little Chinese

SC30 这套沙发很新

zhè tào shāfā hěn xīn
this set sofa very new
This sofa is very new

SC31 我的头疼死了

wǒ de tóu téng sǐ le
1p SBORD head hurt death CRS
I got a bad headache

SC32 我们从美国来

wǒ-men cóng Měiguó lái
1p-PL from USA come
We have come from the USA

SC33 他们意见不同

tā-men yìjian bù tóng
3p-PL idea NEG same
Their ideas are not the same

SC34 这层房租很贵

zhè céng fáng-zū hěn guì
this floor room-rent very expensive
The rent is expensive on this floor

SC35 我伯伯是医生

wǒ bóbo shì yīshēng
1p father's.elder.brother be doctor
My father's elder brother is a doctor

SC36 学生们很高兴

xuéshēng-men hěn gāoxìng
student-PL very happy
The students are happy

SC37 不要在这儿付钱

bú yào zài zhèr fù qián
NEG want at here pay money
I don't want to pay here

SC38 水平提高很多

shuǐpíng tí-gāo hěn duō
level raise-high very much
The level has improved a lot

SC39 老师给我作业

lǎoshī gěi wǒ zuòyè
teacher give 1p homework
The teacher gives me homework

SC40 每天很早起床

měi-tiān hěn zǎo qǐ-chuáng
every-day very early rise-bed
I get out of bed very early every day

Appendix 2: Abbreviations

Abbreviations used in the paper
DA-*t*: The Danish aspirated (affricated) alveolar plosive in simple onset [tˢ]
EN-*t*: The English aspirated alveolar plosive in simple onset as spoken by a native speaker of Danish
L1: Native language
L2: Second language
NLM: Kuhl's Native Language Magnet model
PAM: Best's Perceptual Assimilation Model
SLM: Flege's Speech Learning Model
SC: Standard Chinese (Mandarin Chinese)
VOT: Voice Onset Time, used here to denote the entire duration from the plosive release to the beginning of voicing of a stop or an affricate, including both aspiration and affrication
Y1: First year student at Chinese Studies
Y2: Second year student at Chinese Studies
Y3: Third year student at Chinese Studies

Abbreviations used in the glosses of Danish sentences (Appendix 1)
1sg, 2sg, 3sg: First, second, and third person singular pronouns
ADVZ: Adverbializing suffix
COM: Common gender
DEF: Definite article or suffix
FEM: Feminine gender
INF: Infinitive particle or suffix
GEN: Genitive case
MASC: Masculine gender
NEU: Neutral gender
NOM: Nominative case
OBL: Oblique case
PL: Plural
PRS: Present tense
PST: Past tense

Abbreviations used in glosses of SC sentences (Appendix 1)
1p, 2p, 3p: First, second, and third person
COMPL: Complement particle
CRS: Currently Relevant State particle (see Li & Thompson 1981)
INCL: Inclusive
NEG: Negative particle
PL: Plural
PRF: Perfective aspect particle
SBORD: Subordinating particle