IS THE VOICELESS PALATAL FRICATIVE DISAPPEARING FROM SPOKEN NORWEGIAN?

Wim A. van Dommelen

Department of Language and Literature, NTNU, Trondheim, Norway
wim.van.dommelen@ntnu.no

ABSTRACT

According to impressionistic observations of spoken Norwegian, there is an increasing tendency to pronounce the voiceless palatal fricative /ç/ as its postalveolar counterpart /ʃ/. Through an acoustic analysis of speech material from a large database, this study examines how far this merging phenomenon has proceeded. The database contains recordings of both read and spontaneous speech produced by younger (< 40 yr) and older (> 40 yr) female and male speakers. Auditory evaluation revealed that merging is relatively rare in the database. It appeared to occur as often in spontaneous as in read speech but more often in younger than in older speakers. On the other hand, spectral measures showed that acoustic-phonetic contrasts between canonical productions of /ç/ and /ʃ/ were stronger in the younger generation. No gender-specific effects were observed in /ç/-/ʃ/ contrasts. On the whole, the findings suggest that the merger process is in its infancy at best.

Keywords: Norwegian, palatal, postalveolar, fricatives, merger

1. INTRODUCTION

Norwegian is one of the relatively few languages of the world that have a voiceless palatal fricative /ç/ in their phonological system. Among the 504 languages contained in LAPSyD [10], merely 14 (corresponding to 2.8 %) have this fricative in their sound inventory. Half of them (1.4 %) also have a voiceless postalveolar /ʃ/. According to observations by the general public, younger speakers of Norwegian often pronounce the /ç/ sound as its postalveolar counterpart /ʃ/, thus neutralizing the distinction between word pairs like *kjekk* /çɛk/ ('handsome') and /ʃɛk/ ('check'). The question is whether this tendency indicates the beginning of a general sound change eventually resulting in the disappearance of the /ç/ phoneme from the Norwegian sound system.

A number of non-instrumental investigations have looked into the occurrence of fricative mergers in spoken Norwegian. Varying degrees of merging among younger speakers (14-25 yr) were reported by [8] (1.9 %), [7] (25.5 %), [13] (around 40 %), [17] (41.7 %) and [5] (approx. 50 %). Frequent merging has been documented in urban environments, especially in recordings from speakers with a multilingual background [13]. At the same time, data from [7] indicate that merging frequency may vary between cities. As to gender-specific effects, mixed results have been reported. More merging in female than in male speakers was found in [8, 14], the opposite in [2, 17]. It should be noted that speakers with fully consistent merging behavior seem to be rare [2, 5, 7, 8, 14, 17]. No occurrences of mergers were observed in speakers ≥ 65 yr in [8, 12], >35 yr in [17]. Acoustic analysis of merged 12-year-olds' tokens in [21] showed that neutralization of the phonemic opposition was not completed.

To shed more light on this issue, the present study uses materials from the Norwegian speech database NB Tale [11]. According to preliminary inspection by the author, in the database there was no evidence of /ʃ/ pronounced as [ç] (cf. the reports on that type of merger in [7, 8]). Therefore, the following analysis deals with the occurrences of /ç/ pronounced as [ʃ] and canonically realized /ç/ and /ʃ/. The main research questions were the following:

- How frequent is the merger?
- Is the merger completed as to its acoustic-phonetic properties?
- Is there more evidence of merging in the speech of younger compared to older speakers?
- Is there more evidence of merging among male vs. female speakers?
- Is the merger equally frequent in read vs. spontaneous speech?

2. METHOD

2.1. Terminology

In the following, standard speech phonemes /ç/ and /ʃ/ and their respective pronunciation as [ç] and [ʃ] will be referred to as *canonical* or *distinct*. Canonical realization of /ç/ will be denoted as [/ç/; ç]; merged pronunciation as [/ç/; ʃ]. Accordingly, speakers will be categorized as *distinct* or *merged pronunciation* speakers. Note that a number of members of the latter group also had canonical [/ç/; ç] tokens in their repertoire. Without exception, /ʃ/ was produced as [ʃ] by all speakers of both groups (denoted as [/ʃ/; ʃ]).

2.2. Speech material

Speech materials used for this investigation were taken from the speech database NB Tale, provided by the National Library of Norway [11]. The corpus was collected in 2012 and contains annotated recordings of 240 native speakers of Norwegian. This material is divided according to speakers' dialectal background into 12 subgroups containing 20 speakers each. The 12 dialect regions cover the main four dialect groups: Northern, Central, Western, and Eastern Norwegian. At the time of the recording, speakers were explicitly told that they were free to use speech sounds from their own dialect (e.g. realizing /n/ as [ɳ] or /l/ as [ɽ]). In addition to dialect, speaker age (18-40 yr vs. 40-80 yr) and gender are included as systematic factors in the database. Within each subgroup, there are thus five younger and five older speakers of each gender.

Almost all recordings in the database were made in a sound-treated studio using 48 kHz sampling frequency and 16-bit quantization. All recordings were in stereo using a Sennheiser 2-1-5 headset and a Shure KSM 44 studio microphone. For the present acoustic analysis, the Shure KSM 44 channel was chosen.

2.2.1. Read speech

Recordings in NB Tale comprise 20 sentences per speaker. Sentences were selected from news text corpora and balanced in order to achieve an even distribution between several phonetic and phonemic properties. Only three sentences were produced by all 240 speakers; the remaining ones varied across speakers. Sound files were annotated in X-SAMPA using Praat TextGrids [1].

2.2.2. Spontaneous speech

In addition to read sentences, the NB Tale speech database also contains spontaneous speech recordings. Before reading the sentences, participants were asked to speak for approximately two minutes about a topic of their own choice. According to impressionistic observations, speaking style in the spontaneous speech recordings is as a rule very informal, as witnessed by the occurrence of strong segmental reductions, assimilations and elisions. These recordings were annotated orthographically using Praat TextGrids. The orthographic transcription was used to collect occurrences of /ç/ and /ʃ/. It appeared that in the spontaneous speech of some speakers no relevant /ç/ tokens are contained at all. Moreover, for 11 speakers no spontaneous recordings are available in the speech database.

2.3. Categorization of /ç/

Categorization of /ç/ realizations as [ç] or [ʃ] was based on the results from a formal auditory evaluation. The speech material collected for this evaluation consisted of 1234 realizations of /ç/ present in the database, extracted from the recordings using a parabolic window of 800 ms with the fricative centered. Additionally, intensity differences were leveled out by adjusting tokens to 70 dB. Stimuli contained two presentations of each token separated by a 1-sec silent pause.

Stimuli were presented via two high-quality loudspeakers to individual listeners seated in a sound-treated recording studio. Listeners rated fricative quality using a Likert scale ranging from 1 (clear [ʃ]) to 7 (clear [ç]) presented on a computer screen. The test took 2-3 hours to complete.

A group of seven native speakers of Norwegian (three females, four males) with a background in linguistics participated. They were aged 25-70 yr (mean 38.9 yr) and had no reported hearing problems. All subjects were paid for their participation.

Subsequently, rater data were evaluated as follows. Cases with scale values from 5-7 from a majority (i.e., at least four) of the raters were classified as distinct pronunciation ([/ç/; ç]). Majority judgments with scale values from 1-3 were categorized as mergers ([/ç/; ʃ]). 32 tokens without majority for either of the two alternatives were excluded. For further results, see section 3.1.

2.4. Acoustic analysis and statistical evaluation

Acoustic analysis was carried out using Praat's scripting facility. The first parameter extracted was fricative duration. Further, long-term average spectra (LTAS) with a filter bandwidth of 250 Hz were calculated from the central 40 ms portion of fricatives longer than 80 ms and from the central 50 % of fricatives shorter than 80 ms extracted using a rectangular window shape. Subsequently, various spectral measures were derived from the LTAS. Due to space limitations, this paper will only report on Center of Gravity (CoG, calculated using Praat's query function) and a derived DCT-based measure (see next paragraph and section 3.4).

Following [22] (cf. also [4]) LTAS energy distributions were parameterized using Discrete Cosine Transformation (DCT). In a first step, after conversion of the Hz to the Bark scale, spectral energy values were averaged in 17 critical bands ranging between 8 and 24 Bark ([20]: equation (6)). The 8 Bark lower limit was chosen to avoid artifacts due to occasional occurrence of low frequency disturbances. Subsequently, three DCT-coefficients C_m (m = 2, 3, 4) were calculated with equation (1):

$$C_m = \frac{2}{N} k_m \sum_{n=0}^{N-1} x(n) \cos\left(\frac{(2n+1)(m-1)\pi}{2N}\right)$$

$$k_m = \frac{1}{\sqrt{2}}, m = 1 \quad k_m = 1, m \neq 1 \qquad (1)$$

While coefficient DCT1 characterizes spectral slope, DCT2 and DCT3 represent the spectral envelope's curvature and the contribution of the higher frequencies, respectively (DCT0 encoding mean energy was not calculated).

Statistical analysis of duration and CoG data was performed with the R program package lme4 [15] to calculate Linear Mixed Effects Models (LMEM). Evaluation included the fixed factors phoneme (palatal, postalveolar), speaker age (< 40 yr, > 40 yr), gender (female, male), and dialectal background (12 subgroups). LMEMs included respectively by-speaker random slopes and intercepts for the factor phoneme (section 3.2) and random speaker intercepts (section 3.3). Rejection level was at $\alpha = 0.05$.

3. RESULTS

3.1. Frequency of occurrence

The first research question of this study concerns the distribution of [ç] and [ʃ] represented in the NB Tale database and, specifically, how many occurrences of [ʃ] are cases of phonemic merging. An overview of the number of present observations is given in Table 1. Because the speech material of 15 out of the 240 speakers did not contain any [ç] occurrences, their [ʃ] tokens were excluded from analysis. In this way, the comparison of [ç] and [ʃ] across speakers was balanced. There were 32 *merged pronunciation* speakers (26 among them produced also canonical [/ç/; ç]).

Table 1: Number of fricatives analyzed in read and spontaneous speech. Number of speakers for [/ç/; ç] and [/ʃ/; ʃ]: 225; for [/ç/; ʃ]: 32.

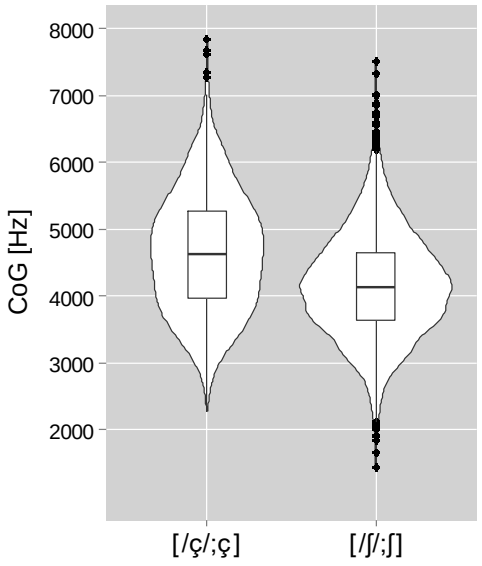
	[/ç/; ç]		[/ç/; ʃ]		[/ʃ/; ʃ]
	n	%	n	%	n
Read	750	93.6	51	6.4	2911
Spont.	374	93.3	27	6.7	728
Total	1124		78		3639

Further, the proportion of phonemic mergers ([/ç/; ʃ]) is relatively small. In read and spontaneous speech, virtually the same proportions of mergers were observed (6.4 % vs. 6.7 %). A chi-square test confirmed that the difference is statistically non-significant ($\chi^2(1) = 0.06$; $p = 0.808$).

3.2. Distinct pronunciation

Inspection of the data from 199 distinct pronunciation speakers revealed that [ç] and [ʃ] had similar mean durations (108 and 110 ms, respectively; according to LMEM calculation $\chi^2(1) = 0.635$; $p = 0.426$). In contrast, mean CoG values were significantly different (4648 Hz and 4161 Hz, respectively; $\chi^2(1) = 64.9$; $p < 0.001$). At the same time, distributions of CoG values were to a relatively large degree overlapping (cf. Fig. 1).

Figure 1: Distribution of CoG values for canonical realizations of /ç/ and /ʃ/ for 199 distinct pronunciation speakers.

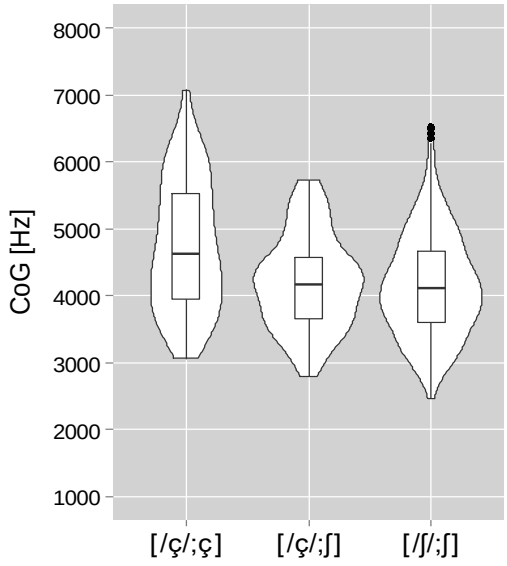


3.3. Merged pronunciation

In this section, we shall investigate the acoustic properties of fricative tokens produced by 32 merged pronunciation speakers (18 females; 14 males). Inspection of their geographical background showed that a relatively large group (12) spoke a Southwest Norwegian dialect, while six had an East Norwegian background. Otherwise, there was no geographical clustering. The majority of the merged pronunciation speakers (29) were younger than 40 years.

The speech material produced by the merged pronunciation speakers contained in total 151 /ç/ occurrences, 73 of which were realized as [ç], 78 as [ʃ]. In the following, we will focus on two questions. The first one is to which degree merged pronunciation speakers' [/ç/; ç] tokens were acoustically different from their [/ʃ/; ʃ] tokens. Inspection of the data showed that mean durations did not differ significantly (113 ms vs. 107; $\chi^2(1) = 1.792$; $p = 0.181$). In contrast, position of CoG on the frequency scale was significantly different (4735 Hz vs. 4150 Hz; $\chi^2(1) = 47.908$; $p < 0.001$; cf. Fig. 2).

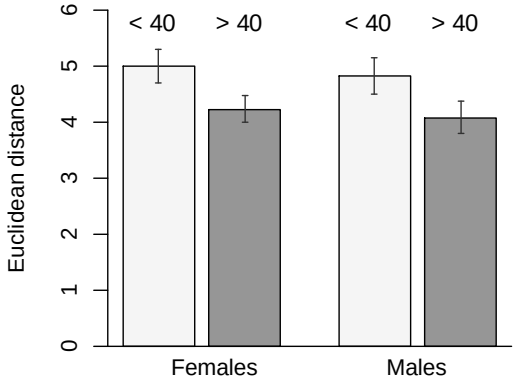
Figure 2: Distribution of CoG values for canonical realization of /ç/ ([/ç/; ç]), /ç/ realized as [ʃ] ([/ç/; ʃ]), and [/ʃ/; ʃ] for 32 merged pronunciation speakers.



The second point of interest is whether merged /ç/ ([/ç/; ʃ]) fully overlapped with [/ʃ/; ʃ]. Data showed that, again, duration does not play a role for the distinction (117 ms for the former, 107 ms for the latter; $\chi^2(1) = 3.357$; $p = 0.067$). Additionally, CoG values were close (4203 Hz vs. 4150 Hz; $\chi^2(1) = 0.001$; $p = 0.981$).

3.4. Speaker age and gender

Figure 3: Mean Euclidean distance in DCT1xDCT2xDCT3 space between [/ç/; ç] and [/ʃ/; ʃ] for 199 distinct pronunciation speakers (100 f; 99 m) aged under and over 40 yr, respectively. Error bars indicate ± 1 standard error.



To investigate the effects of age and gender, spectral measures DCT1 to DCT3 were used. Calculations were performed as follows. For each speaker, mean values of DCT1 to DCT3 were computed for [/ç/; ç] and [/ʃ/; ʃ] tokens separately. Next, Euclidean distances separating them in the DCT1xDCT2xDCT3 space were calculated. In this way, for each speaker one Euclidean distance value was obtained quantifying the contrast between the two types of fricatives.

Figure 3 depicts the results for female and male speakers, separated by age category. Acoustic contrasts were larger for younger than for older speakers (mean values of 5.09 and 4.18, respectively) and, to a lesser extent, larger for females than males (4.68 vs. 4.47). According to a two-way ANOVA with speaker age and gender as factors, only the former was statistically significant ($F(1, 195) = 8.93$; $p = 0.003$; gender: $F < 1$). There was no significant age x gender interaction ($F < 1$).

4. DISCUSSION AND CONCLUSION

In this investigation, the overwhelming majority of /ç/ occurrences were actually produced with a palatal place of articulation. At the same time, however, spectral differences between canonically realized /ç/ and /ʃ/ were relatively small (cf. [18, 19]). In view of the articulatory similarity and overlap as well as perceptual similarity of the two phonemes, in [19] conditions for a merger were considered to be good. However, in the group of distinct pronunciation speakers, acoustic-phonetic contrasts between the palatal and the postalveolar fricative appeared to be stronger in the younger than in the older generation. This may be seen as evidence against a beginning disappearance of the /ç/ phoneme from the Norwegian sound system in spite of merging being more frequent among younger speakers.

Further, only six merged pronunciation speakers were consistent in their behavior (cf. similar results reported in [2, 5, 7, 8, 14, 17]).

Previous studies have often found female speakers to make clearer contrasts between sound categories than males [3, 6, 9, 16, 23]. In this investigation, however, Euclidean distances in the DCT1xDCT2xDCT3 space were not gender-dependent. All above-mentioned studies involved /s/ vs. /ʃ/ contrasts. Since gender-marking seems to be restricted to the alveolar fricative, this can maybe explain the absence of gender-specific effects for the present /ç/-/ʃ/ contrast.

In conclusion, the results of this study have shown that pronunciation of the Norwegian palatal fricative /ç/ as [ʃ] is relatively rare, at least in formal recording situations. There are no strong indications of increasing use of the merger. Although mergers were observed more often in younger vs. older speakers, in the younger generation acoustic-phonetic contrasts between canonical productions of /ç/ and /ʃ/ were stronger. It seems safe to conclude that at present the position of the Norwegian /ç/ is not endangered.

6. REFERENCES

[1] Boersma, P., Weenink, D. 2015. Praat: Doing phonetics by computer (version 6.0). http://www.praat.org/ (accessed 29 October 2015).

[2] Dalbakken, Liv O. 1996. *Distinksjonen kje/sje i lydendringsperspektiv. En empirisk undersøkelse av barns og unges beherskelse av distinksjonen kje/sje i Trondheim* [The *kje/sje* distinction in the perspective of sound change. An empirical investigation of children's and young people's mastery of the *kje/sje* distinction in Trondheim]. Graduate thesis, University of Oslo.

[3] Fox, R. A., Nissen, S. L. 2005. Sex-related acoustic changes in voiceless English fricatives. *Journal of Speech, Language, and Hearing Research* 48, 753–765.

[4] Harrington, J. 2010. *The phonetic analysis of speech corpora*. Malden, MA : Wiley-Blackwell.

[5] Hårstad, S. 2010. *Unge språkbrukere i gammel by: en sosiolingvistisk studie av ungdoms talemål i Trondheim* [Young language users in an old city: a sociolinguistic study of youth language in Trondheim]. Ph.D. thesis, Norwegian University of Science and Technology.

[6] Holliday, J. J., Reidy, P. F., Beckman, M. E., Edwards, J. 2015. Quantifying the robustness of the English sibilant fricative contrast in children. *Journal of Speech, Language, and Hearing Research* 58, 622–637.

[7] Jacobsen, A. S. 2015. *"Vi ungdomma sir jo sånn /ʃ/ i stedet for /ç/". Ei sosiolingvistisk sammenligning av fonemkutviklinga av /ʃ/ og /ç/ i Tromsø og Trondheim* ["We youngsters say such a /ʃ/ instead of /ç/, you know". A sociolinguistic comparison of the phoneme development of /ʃ/ and /ç/ in Tromsø and Trondheim]. Master's thesis, UiT The Arctic University of Norway.

[8] Johannesen, S. H. 1983. *Om 'skjendisar' og 'chipsreiarar'. Bruken av sje-lyd og kje-lyd i bergensmålet* [About 'skjendisar' and 'chipsreiarar'. The use of the 'sje' sound and the 'kje' sound in the dialect of Bergen]. In: Sandøy, H. (ed.), *Talemål i Bergen* 1. Bergen: Nordisk institutt, University of Bergen, 5–28.

[9] Jongman, A., Wayland, R., Wong, S. 2000. Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am*. 134, 1252–1263.

[10] LAPSyD - Lyon-Albuquerque Phonological Systems Database. http://www.lapsyd.ddl.ish-lyon.cnrs.fr/ (accessed 26 March 2018).

[11] NB Tale – en grunnleggende akustisk fonetisk taledatabase for norsk [NB Speech – a fundamental acoustic-phonetic speech database for Norwegian]. http://www.nb.no/sprakbanken/show?serial =sbr-31&lang=nb (accessed 11 June 2015).

[12] Nødtvedt, A. F. 1991. *Molde bymål. Språklege variasjonar hos eldre og yngre brukarar* [The urban dialect of Molde. Variations in the language of older and younger users]. Graduate thesis, University of Trondheim.

[13] Opsahl, T., Røyneland, U. 2009. *Osloungdom – født på solsiden eller i skyggen av standardtalemålet?* [Oslo youth – born on the sunny side or in the shadow of the standard language?] *Norsk lingvistisk tidsskrift* 27, 95–120.

[14] Papazian, E. 1994. *Om sje-lyden i norsk, og ombyttinga av den med kje-lyden* [About the 'sje' sound in Norwegian, and its substitution with the 'kje' sound]. *Norskrift* 83, 1–105.

[15] R Core Team (2016). R: A language and environment for statistical computing. Retrieved from https://ww.R-project.org/

[16] Romeo, R., Hazan, V., Pettinato, M. 2013. Developmental and gender-related trends of intra-talker variability in consonant production. *J. Acoust. Soc. Am*. 134, 3781–3792.

[17] Rød, H. M. 2014. *Molde bymål. Ein kvantitativ korrelasjonsstudie i verkeleg tid* [The urban dialect of Molde. A quantitative correlation study in real time]. Master's thesis, University of Bergen.

[18] Simonsen, H. G., Moen, I. 2000. */ʃæ:re ʃiʃti/: om skillet mellom /ʃ/ og /ç/ i et fonetisk perspektiv* [/ʃæ:re ʃiʃti/ 'Dear Kirsti': about the distinction between /ʃ/ and /ç/ from a phonetic perspective]. In: Andersen, Ø., Fløttum, K., Kinn, T. (eds.), *Menneske, språk og fellesskap. Festkrift til Kirsti Koch Christensen på 60-årsdagen 1. desember 2000*. Oslo: Novus Forlag, 175–193.

[19] Simonsen, H. G., Moen, I. 2004. On the distinction between Norwegian /ʃ/ and /ç/ from a phonetic perspective. *Clinical Linguistics and Phonetics, 18* (6–8), 605–620.

[20] Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *J. Acoust. Soc. Am*. 88, 97–100.

[21] van Dommelen, W. A. 2003. An acoustic analysis of Norwegian /ç/ and /ʃ/ as spoken by young people. *J. Int. Phon. Ass.* 33, 131–141.

[22] Watson, C. I., Harrington, J. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *J. Acoust. Soc. Am*. 106, 458–468.

[23] Weirich, M., Simpson, A. P. 2015. Gender-specific differences in sibilant contrast realizations in English and German. *Proc. 18th ICPhS* Glasgow, UK, Paper number 0261.