Perceptual assimilation of English dental fricatives by native speakers of European French

Michael D. Tyler[1], Eléonore Clot[2], Marie-Sophie Villain--Bailly[2], Chotiga Pattamadilok[2]

[1]School of Social Sciences and Psychology, Western Sydney University, Australia
[2]Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France
m.tyler@westernsydney.edu.au, chotiga.pattamadilok@lpl-aix.fr

ABSTRACT

The phonetic characteristics of French-accented speech suggest that French native speakers often have difficulty producing dental fricatives in English. However, there is a surprising lack of empirical research on perception of those consonants. Canadian French speakers appear to assimilate /θ/ to /t/ and /ð/ to /d/, but loanword evidence suggests that European French speakers should assimilate them to /s/ and /z/, respectively. To test this, 151 native European French listeners categorised and rated the goodness-of-fit of English /θ, f, s, t, ð, v, z, d/ to French phonological categories. /θ/ was categorised as /f/, whereas /ð/ was uncategorised, with responses divided between /v/ and /z/. The remaining consonants were categorised as their corresponding French categories, with /θ/ rated as a poorer French /f/ than /f/. While the majority of individual participants categorised the dental fricatives as /f, v/, there were small subsets of participants who categorised them as /s, z/.

Keywords: Cross-language speech perception, perceptual assimilation, dental fricatives, French.

1. INTRODUCTION

What is ze meaning of zis, Dumbly-dorr?
Madame Maxime,
Harry Potter and the Goblet of Fire

English native speakers have the impression that a hallmark of European-French-accented English is substitution of /s/ and /z/ for /θ/ and /ð/. What little evidence there is from loanword phonology seems to concur, with observations of *Granny Smith* /gʁanismis/ and *Big Brother* /bigbʁɔzœʁ/ [12] (and, more recently, *Bluetooth* /blutus/ [16] and *The Voice* /zəvɔis/ [17]). There is some evidence for substitution with /t/, as in *macarthysme* and *cathode* [4], although this may be orthographically based, as *th* in French is usually pronounced as /t/.

Studies on production of English dental fricatives by French speakers are scarce. There is general agreement that Canadian French speakers produce /θ/ and /ð/ as /t/ and /d/, respectively [5, 11-13], and that European French speakers produce them as /s/ and /z/ [5, 10, 12]. However, a small-scale production study of European French speakers showed that most speakers produced the dental fricatives correctly most of the time in reading and imitation tasks [15]. When they were produced incorrectly, they were produced as /f, v/ rather than /s, z/.

This apparent mismatch between loanword evidence and production raises the question of how English dental fricatives are perceived by European French listeners. Canadian French listeners perceive words containing /θ/ as homophonous with minimal pairs containing /t/ (e.g., *three*/*tree*), in line with their productions [13]. Surprisingly, we have not found any systematic investigation of the perception of dental fricatives by speakers of European French in the peer-reviewed literature. The aim of the present study is to investigate how English dental fricatives are perceived relative to native European French consonant categories. We test a large group of over 150 European French listeners to analyse individual differences in perception of /θ/ and /ð/.

As a framework, we use the Perceptual Assimilation Model (PAM) [1], which outlines different ways that non-native phones might be assimilated to the native phonological space. Non-native phones may be *categorised* as a native phonological category (with a goodness-of-fit ranging from poor to excellent) or *uncategorised* if it is perceived as speech but as unlike any single native category. Uncategorised phones can be further subdivided into *focalised* (weakly consistent with a single native category), *clustered* (weakly consistent with more than one native category), or *dispersed* (not consistent with any native category) [8]. Non-native phones that are perceived as non-speech (e.g., click consonants by English listeners [2]) are *non-assimilable*. PAM makes predictions about discrimination of pairs of minimally contrasting phones based on these perceptual assimilations.

European French native speakers were asked to categorise English consonants in terms of their native French consonant categories and rate the goodness of fit to the chosen category. The target consonants were /θ, ð/ and the English consonants that they are most likely to be confused with: /f, s, t, v, z, d/. We took a whole-system approach [6, 7] to categorisation by giving participants the opportunity to select from all possible French consonants. It should be noted that all French school students now learn English at school, so it is not possible to find university-aged students who are naïve to English. The data presented here are part of a larger project where categorisation in French and English will be compared to discrimination accuracy. This paper focuses on perceptual assimilation to French only, to examine how experience with French as a native language shapes the perception of English dental fricatives and the variability of perceptual assimilation across individuals.

If perception of English dental fricatives matches predictions of loanword phonology then /θ/ and /ð/ should be assimilated to /s/ and /z/, respectively. If European French speakers' assimilations are similar to Canadian French speakers' then they should instead be assimilated to /t/ and /d/. Finally, if the observations of [15] in production tasks hold true for perception, then assimilation should be to /f/ and /v/.

2. METHOD

2.1. Participants

The participants were 151 native French speakers recruited from a university sample (106 female, 136 right-handers, M_{age} = 22.2 years, *SD* = 3.5, range: 18-46). Around half of the participants grew up in the south of France, with the remainder from a range of French regional areas. All of the participants learned English and at least one other language at school, usually Spanish or German. The mean age of commencement of English study was 9.1 years (*SD* = 2.0, range: 3-13).

2.2. Stimuli

A phonetically trained female native speaker of Australian English produced tokens of 23 English consonants in a consonant + /a/ frame (/p, b, t, d, k, g, m, n, f, v, θ, ð, s, z, ʃ, ʒ, h, tʃ, dʒ, w, r, l, j/). The four best tokens of the target consonants /θ, ð, f, v, s, z, t, d/ were selected on the basis of auditory and visual inspection of the waveform and spectrogram. The tokens were selected on that basis, rather than on the basis of expected acoustic characteristics, to allow natural acoustic variability among tokens. The best single tokens of the each of the remaining filler consonants were selected in the same way.

The stimuli were recorded in a sound-attenuated booth using a Beyerdynamic TG H55c microphone and an Edirol UA-25EX USB audio capture device. The audio was recorded using Audacity at a 44.1 kHz sampling rate with 16-bit resolution. The session recording was high-pass filtered at 70 Hz to remove low-frequency rumble and correct for the DC component. The duration of the vowel was truncated to 80 ms, and a 5 ms ramp was applied, to minimise the influence of the vowel on categorisation [9]. The acoustic characteristics of the final target stimuli are presented in Table 1. All acoustic analyses and stimulus editing were performed using Praat [3]. The sound files were downsampled to 22.05 kHz and high-pass filtered from 300 Hz to analyse the spectral moments, which were calculated on the basis of a 25.6-ms Hamming-windowed slice centered at the 50% duration point of frication. Power was set to 1.

Table 1: Acoustic measurements of critical stimuli. Stim = Stimulus, Dur = Duration (ms), Int = Intensity (dB), COG = Centre of Gravity (Hz), SD = Standard Deviation (Hz), Skew = Skewness, Kurt = Kurtosis.

Stim	Dur	Int	COG	SD	Skew	Kurt
θa1	141	52	6547	3028	-0.177	-1.150
θa2	118	45	7276	3293	-0.705	-0.879
θa3	127	48	6370	3251	-0.292	-1.159
θa4	113	47	6344	3349	-0.324	-1.246
fa1	133	58	6819	2934	-0.485	-0.906
fa2	146	54	6596	3145	-0.274	-1.215
fa3	153	58	7003	2959	-0.574	-0.887
fa4	137	53	6802	3090	-0.427	-1.091
sa1	146	66	8385	2226	-1.388	1.850
sa2	146	60	7627	2403	-1.153	1.089
sa3	139	65	8663	1957	-1.780	3.733
sa4	97	66	8710	1802	-2.057	5.803
ta1	85	60				
ta2	70	59				
ta3	55	56				
ta4	63	57				
ða1	130	66	5305	3505	0.018	-1.431
ða2	123	66	5690	3287	-0.081	-1.278
ða3	132	64	5825	3362	0.005	-1.306
ða4	124	63	6072	3042	-0.105	-1.197
va1	135	68	5754	3562	-0.139	-1.406
va2	140	64	6571	3527	-0.445	-1.216
va3	148	64	5443	3499	0.030	-1.442
va4	107	63	5648	3288	-0.046	-1.248
za1	141	66	8278	2130	-1.747	3.718
za2	155	64	8428	2286	-1.682	2.747
za3	137	63	7902	2640	-1.497	1.714
za4	138	64	8063	2160	-1.718	3.475
da1	68	50				
da2	131	62				
da3	124	58				
da4	110	60				

The response categories were 22 orthographic labels, including all attested French consonants (including those only occurring in loanwords), with keywords provided for potentially ambiguous letters: B, CH *chat*, D, DJ *jazz*, F, G *gant*, GN *gnôle*, HU *huit*, J, K, L, M, N, P, R, S, T, TCH, V, W *wapiti*, Y *yaourt*, Z. These corresponded to the consonant categories /b, ʃ, d, dʒ, f, g, ɲ, ɥ, ʒ, l, m, n, p, ʁ, s, t, tʃ, v, w, j, z/. Stimulus presentation and response collection in the experimental tasks were controlled by E-Prime software.

2.2. Procedure

Participants were tested individually or in pairs. The session consisted of six discrimination tasks followed by two categorisation tasks. The focus of this paper is on the first categorisation task.

The experimenter first familiarised the participants with a printed version of the response grid. The participants were told that they would hear a syllable through headphones and they should pay attention to the consonants at the beginning of the syllable. They were instructed to select from among the French consonant categories provided, even if the consonant resembled one from another language that they knew. If they were unsure, they should guess. After selecting the consonant, they rated the consonant's goodness of fit to the selected category on a scale of 1 to 7, where 1 was a poor fit (*très peu fidèle*), 4 was medium (*correspond moyennement*) and 7 was perfect (*correspond parfaitement*).

On each trial participants heard a single token through headphones, and a grid showing the response labels in alphabetical order. After clicking on a response, the screen was immediately replaced with a rating screen. No feedback was provided. Participants were given 4 s to categorise the stimulus, and 3 s to provide a rating. If the participant timed out on either the labelling or the rating, the trial was randomly reinserted into the trial list. There were three blocks of randomised trials and each target consonant token was presented once per block. The filler consonants were included so that the participants would not rely on a small number of response categories, and to check for vigilance and understanding of the task. The fillers were presented twice each across the three blocks, interspersed among the target consonants, with the constraint that the same filler consonant was not presented twice in the same block. There were 126 trials in total (8 target consonants x 4 tokens x 3 blocks + 15 filler consonants x 2 presentations). The intertrial interval was 1 s and the task took around 15 minutes to complete.

3. RESULTS

To test the reliability of the categorisations across the group, the mean number of choices for each label × stimulus combination was tested against a chance score of 1/22 using a one-sample *t*-test. For target consonants, the mean percent categorisation and mean goodness ratings for above-chance responses are presented in Table 2.

Table 2: Percent categorisations to French consonant categories and mean goodness ratings out of 7 (in parentheses) for each target stimulus category (Stim. Cat.). Labels not selected above a chance level of responding were omitted from the table.

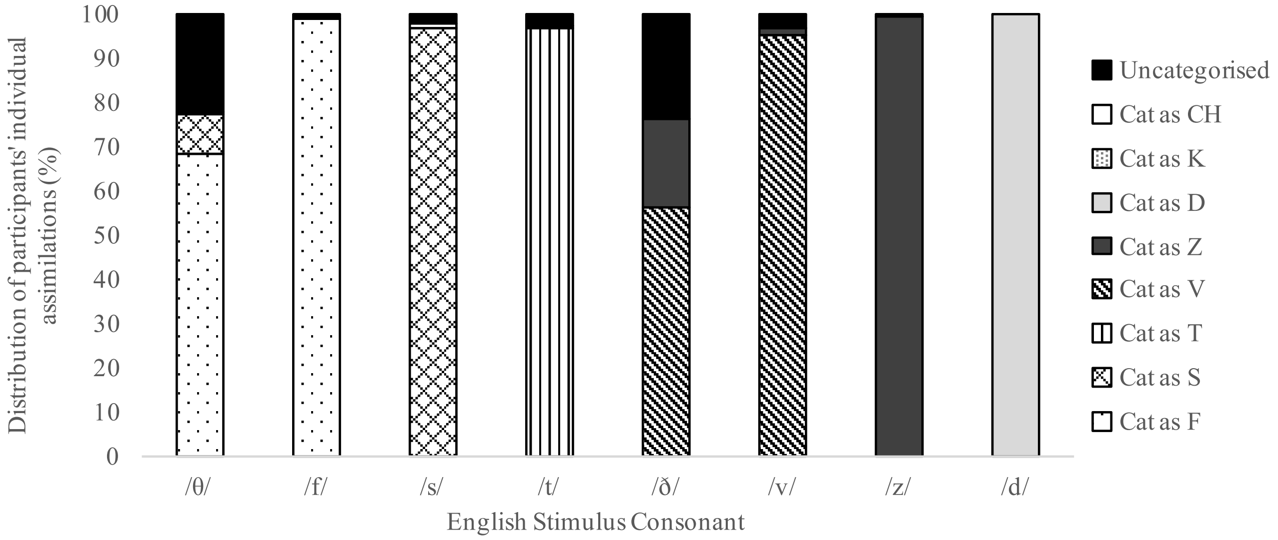
Stim.	French Category Label					
Cat.	F	S	T	V	Z	D
/θ/	76 (4.7)	19 (3.4)				
/f/	98 (6.1)					
/s/		97 (6.0)				
/t/			97 (5.7)			
/ð/				67 (4.9)	31 (3.8)	
/v/				96 (6.1)		
/z/					99 (6.4)	
/d/						99 (5.9)

Using a 70% threshold for categorisation [for discussions on the appropriate use of thresholds, see 6, 8], /f, s, t, v, z, d/ were each categorised as the corresponding French consonant. /θ/ was categorised as F with a goodness-of-fit falling in the medium range. There was also a small but significant selection as a medium-poor version of S. The difference in mean goodness rating as F for /θ/ versus /f/ was significant, *Welch's t*(225.56) = 10.33), $p < .001$. /ð/ was uncategorised/clustered, with responses divided between medium-goodness V and medium-poor Z.

With an analysis by group it is not possible to determine whether the split responses are due to many participants perceiving the English fricatives as similar to more than one French category, or to subsets of participants perceiving them differently from each other. To investigate this, we conducted a second analysis by individual participant. A binomial test was used to establish whether a given label was selected above chance. Again, labels selected above 70% for a given auditory stimulus category were deemed to be categorised. The distribution of individual assimilation types is presented in Figure 1.

Unsurprisingly, individual participants were consistent in their perceptual assimilations of /f, s, t, v, z, d/. The majority of participants categorised /θ/ as F (68%) with only 9% of participants categorising /θ/ as S. Around 58% of participants categorised /ð/ as V, with 20% categorising it as Z. Around a quarter of participants assimilated /θ/ and /ð/ as uncategorised (22% and 24%, respectively).

Figure 1: Distributions of individual participants' perceptual assimilations for each English stimulus consonant. "Cat as" denotes "Categorised as".



4. DISCUSSION

The analysis at the group level showed that /θ/ was assimilated as F. /ð/ was uncategorised/clustered, with responses divided between V and Z, although the modal response was V and it was just below the threshold of 70%. This finding suggested that there may be subgroups of participants who perceptually assimilate the dental fricatives in different ways. An analysis of individual participant assimilations showed that the majority assimilated the dental fricatives as F and V, but with clear subgroups who assimilated them as S and Z. A minority of participants assimilated them as uncategorised.

It is clear from these results that European French listeners do not perceptually assimilate English dental fricatives in the same way as Canadian French listeners (i.e., /t, d/), whose perception may be influenced by the assibilation of their dental /t, d/ to [t^s/d^z] in certain phonetic contexts [14]. English speakers' characterisations of a French accent in production (i.e., /s, z/) do not match with French listeners' perception as predominantly /f, z/. The findings of [15] suggest that dental fricatives are at least sometimes produced as /f, v/, but in loanword phonology production of dental fricatives is /s, z/. This opens up the intriguing possibility that some French native speakers may perceive /θ, ð/ as /f, v/ but produce them as /s, z/.

Direct investigation of a perception-production mismatch would require a targeted experiment, but there are some suggestions in the literature for how such a mismatch might occur. On the basis of a cross-language survey of loanwords, [12] suggested that the adaption of /θ, ð/ to /t, d/ is a phonological adaptation, through [+continuant] delinking, whereas adaptation to /f, v/ is the result of naïve phonetic approximation, or "faulty perception". Adaptation to /s, z/ was thought to be due to difficulties articulating the dental fricatives, with /s/ and /z/ surfacing as production-based phonetic approximations. However, loanword phonology may be influenced by orthography or reflect phonology prior to diachronic change. Participants would need be tested on both perception and production for direct evidence of a mismatch. [5] suggested speakers are aware that the labiodental is the incorrect pronunciation, because it is visually distinct, and opt for an alternative coronal consonant instead. She predicted that French learners acquiring English with auditory materials alone may be likely to produce labiodentals more frequently than those exposed to auditory-visual speech.

In spite of English speakers' impressions that a French accent is characterised by substitution of /θ, ð/ with /s, z/, these results have shown that the majority of native European French listeners perceptual assimilate them as slightly deviant versions of their native labiodental categories. The finding that subgroups of listeners perceive them as /s/ and /z/ warrants further investigation. These data are part of a larger project examining the acquisition of dental fricatives. Future analyses will consider both L1 and L2 categorisation, the influence of perceptual assimilation on discrimination, and whether the individual differences in perceptual assimilation correspond to differences in discrimination.

5. ACKNOWLEDGEMENTS

The research was supported by grants ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

6. REFERENCES

[1] 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*. Baltimore: York Press, 171-204.

[2] Best, C. T., McRoberts, G.W., Sithole, N. M. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance* 14, 345-360.

[3] Boersma, P. 2001. Praat: a system for doing phonetics by computer. *Glot International* 5, 341-345.

[4] Bogaards, P. 2008. *On ne parle pas franglais: La langue française face à l'anglais*. Brussels: De Boeck.

[5] Brannen, K. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics* 47, 1-46.

[6] Bundgaard-Nielsen, R. L., Best, C. T., Tyler, M. D. 2011. Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition* 33, 433-461.

[7] Bundgaard-Nielsen, R. L., Best, C. T., Tyler, M. D. 2011. Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics* 32, 51-67.

[8] Faris, M. M., Best. C. T, Tyler M. D. 2018. Discrimination of uncategorised non-native vowel contrasts is modulated by perceived overlap with native phonological categories. *Journal of Phonetics* 70, 1-19.

[9] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107, 2711-2724.

[10] Hancin-Bhatt, B. 1994. Segment transfer: A consequence of a dynamic system. *Second Language Research* 10, 241-269.

[11] Jamieson, D. G., Morosan, D. E. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/contrast by francophones. *Perception & Psychophysics* 40, 205-215.

[12] Paradis, C., LaCharité, D. 2012. The influence of attitude on the treatment of interdentals in loanwords: Ill-performed importations. *Catalan Journal of Linguistics* 11, 97-126.

[13] Trofimovich, P., John, P. 2011. When three equals tree: Examining the nature of phonological entries in L2 lexicons of Quebec speakers of English. In: Trofimovich, P., McDonough, K. (eds.), *Applying priming methods to L2 learning, teaching and research: Insights from Psycholinguistics*. Philadelphia, PA: John Benjamins, 105-129.

[14] Walker, D. C. 1984. *The Pronunciation of Canadian French*. Ottawa: University of Ottawa Press

[15] Wenk, B. 1979. Articulatory setting and de-fossilization. *Interlanguage Studies Bulletin* 4, 202-220.

[16] https://youtu.be/zf8zb6LB0SM?t=8

[17] https://youtu.be/-CkRGHsrmYs?t=47