PERCEPTUAL ASSIMILATION AND GRADED DISCRIMINATION AS PREDICTORS OF IDENTIFICATION ACCURACY FOR LEARNERS DIFFERING IN L2 EXPERIENCE: THE CASE OF DANISH LISTENERS' PERCEPTION OF ENGLISH INITIAL FRICATIVES

Ocke-Schwen Bohn and Anne Aarhøj Ellegaard

Department of English, Aarhus University, Denmark
ocke.bohn@cc.au.dk, anneandersen3@msn.com

ABSTRACT

The present study had two aims: To examine how well two measures of perceived similarity of L2 sounds (perceptual assimilation and graded discrimination) predict identification accuracy for nonnative consonants, and to examine how L2 experience affects these measures and identification accuracy. Two groups of native Danish listeners differing in English language experience participated. Experiment 1 examined the perceptual assimilation of English initial fricatives to Danish categories, Experiment 2 examined how similar the participants perceived English fricative pairs to be. The results from these experiments generated predictions for the identification accuracy of English fricatives, which were examined in Experiment 3.

Results revealed, in general, the expected effect of L2 experience on all three perceptual measures. Perceived dissimilarity and perceptual assimilation predicted identification accuracy equally well, with one important exception where both measures failed. We discuss whether dissimilarity ratings could replace perceptual assimilation tasks in studies aimed at predicting L2 identification accuracy.

Keywords: L2 speech perception, perceptual assimilation, effect of L2 experience

1. INTRODUCTION

The two most widely used models of L2 speech, the Speech Learning Model (SLM) [6] and the Perceptual Assimilation Model for L2 (PAM-L2) [2], predict that perception and production problems for nonnative sounds are due to how L2 sounds perceptually map onto native categories. These models also predict that L2 experience will affect how L2 sounds are perceptually assimilated to L1 categories, which in turn affects their learnability.

SLM and the precursor of PAM-L2, PAM [1], were originally designed with different aims: SLM predicts ultimate production accuracy of individual L2 sounds, and PAM predicts discrimination levels for nonnative contrasts. However, both models have also been extended to test predictions for other aspects of L2 speech learning. For example, predictions inspired by SLM were examined in studies of L2 identification [5] and discrimination [8], and PAM predictions were tested in studies of L2 identification [4] and production accuracy [7].

The aims of the present study were twofold: First, we wanted to examine whether two measures of perceived similarity of L2 sounds, perceptual assimilation and within-L2 graded discrimination, can successfully predict L2 identification. Perceptual assimilation patterns have frequently been used to predict L2 discrimination and production, and they have also been used to predict the identification of stimuli from synthetic continua [4], but they have rarely been used to predict the identification of a whole set or subset of L2 segments [9]. Second, this study examines the effect of L2 experience on the perceptual assimilation, the graded discrimination, and the identification of L2 sounds. We asked not only whether identification patterns change as a function of L2 experience, but also whether one of the measures, perceptual assimilation or graded discrimination, is more sensitive to L2 experience.

We pursued these questions by examining L1 Danish listeners' perception of English fricatives. An earlier study [9] had revealed that the identification accuracy of Danish listeners for English syllable-initial consonants was quite high for approximants (% correct rates 84.2 to 100) and stops (% correct rates 97.5 to 100), but comparatively low for fricatives (% correct rates 68.1 to 89.2). Table 1 compares the inventories of initial fricatives in English and Danish. English has eight fricatives at four places of articulation with voicing contrasts at all four places. Danish has only three fricatives, all of which are voiceless. Table 1 also lists the Danish labiodental approximant [ʋ], but not the Danish dental approximant [ð̞], which only occurs in postvocalic position. Table 1 shows that English and Danish have [f, s] in common. English has a voiced labiodental fricative, whereas Danish has a voiced labiodental approximant, and the postalveolar voiceless fricative is produced with a larger anterior cavity in English than in Danish, due to lip rounding and a more

retracted tongue position for English [ʃ] than for Danish [ɕ].

Table 1: Comparison of English and Danish fricative inventories (initial position).

English	f	v		θ	ð	s	z	ʃ		ʒ
Danish	f		ʊ			s			ɕ	

2. GENERAL METHODS

We report experiments in which L1 Danish listeners were presented with [Cɑ] syllables in which the initial consonant was one of the eight English fricatives listed in Table 1, or one of the English approximants [w, j, r, ɬ] or affricates [ʧ, ʤ]. Only the results from the perception of the fricatives will be reported here.

The stimuli were taken from a corpus of English [Cɑ] syllables made available by [10]. We selected three tokens each from two male L1 American English speakers in which [C] was one of the 14 EN consonants [f, v, θ, ð, s, z, ʃ, ʒ, w, j, r, l, ʧ, ʤ].

Two groups of L1 Danish listeners participated: An "inexperienced" group who had not spent more than one month in an English speaking country (n = 16, 13f, *M* age: 28.1 (SD=2.6), and an "experienced" group which had spent a mean of 10.4 months (SD=19.4) in an English-speaking country (n = 12, 9f, *M* age: 28.5 (SD=3.5).

The experiments reported below were conducted either in a sound attenuated booth at Aarhus University or in a quiet environment at participants' homes. Experiments were run using *praat* [3] from a PC laptop with high quality headphones.

3. EXPERIMENT 1

3.1 Methods

Experiment 1 examined the perceptual assimilation of the eight English fricatives to Danish. A screen display presented listeners with 11 orthographic Danish response alternatives *f, v, s, sj, l, j, r, tj, dj, j, w* plus an *Intet match* ("no fit") button. The alternatives correspond unambiguously to the Danish consonants /f, v, s, ɕ, l, j, r, tj, dj, w, 1/. Participants were instructed to first click on the Danish letter which was the best match for the English consonant, and then rate the goodness of fit of the English stimulus to the chosen Danish category on a 9-point Likert scale, ranging from 0 (*dårlig* "bad") to 8 (*god* "good"). Participants responded to 84 trials (14 English consonants x 2 talkers x 3 tokens) with an ITI of 0.5 seconds.

3.2 Results and discussion

Table 2 presents the results of the perceptual assimilation task for inexperienced (IN) and experienced (EX) Danish listeners. Following [8], the categorical assimilations and the graded goodness ratings are combined into a fit index, which is derived from multiplying the proportion of assimilations to a particular Danish category with the goodness of fit rating for this assimilation. Thus, the perfect fit would be quantified as 8.0 (proportion of assimilation = 1.0 x highest goodness rating = 8).

Table 2: Mean fit indices (see text) for English [Ca] tokens. Top of cell: Inexperienced listeners; bottom of cell: Experienced listeners. "No fit" responses are given in percentages.

English stimuli	Danish response (and rating)						No fit
	f	ʊ	s	ɕ	dj	w	
f	7.2 6.8						
v		3.8 5.5				2.3	
θ	6.1 3.3						40.3%
ð					0.9		33.0% 93.1%
s			6.5 6.5				
z			5.3 2.3				
ʃ				6.5 5.2			
ʒ				2.8 3.5	1.0		

Table 2 shows that IN and EX differ considerably in their assimilation patterns for [θ, v, ð, z, ʒ] but not much for [f, s, ʃ]. Both groups are expected to identify [f, s, ʃ] highly accurately because they fit Danish counterparts very well. EX assimilate English [v] uniquely and with a good fit to Danish /ʊ/, but IN assimilate [v] to both Danish /ʊ/ and /w/, which should be reflected in lower identification accuracy for IN than EX. A clear difference in identification accuracy is expected for [θ], which IN perceive to have nearly the same fit to Danish /f/ as English [f], suggesting low identification accuracy, whereas EX perceive a poor fit to Danish /f/ or no fit at all. The assimilations for English [ð] suggest that both IN and EX will identify [ð] quite accurately because it does not fit well (IN) or at all (EX) with any Danish category. The difference between the perceived fit of English [z] to Danish /s/ between IN (good fit) and EX (poor fit) suggests EX will identify [z] more

accurately than IN. Surprisingly, IN perceive a poorer fit of English [ʒ] to Danish /ɐ/ than EX, which could suggest higher identification accuracy for IN than EX.

4. EXPERIMENT 2

4.1 Methods

Experiment 2 examined the graded discrimination of selected English consonant pairs. Listeners rated the dissimilarity of the pairs of interest, [f-θ], [v-f], [s-z], [ʃ-ʒ], [v-w], presented six times each, as well as [w-ʃ], [w-ʒ], w-θ], and [w-ð] and nine type-identical pairs (e.g., [f-f]), each presented only once. The [Ca] tokens in each pair were always from different talkers and presented with in ISI of 0.8 seconds. Participants rated the degree of perceived similarity of each randomly presented pair on a 9-point Likert scale ranging from 0 (*ens* "identical") to 8 (*meget forskellige* "very different").

4.2 Results and discussion

Table 3 lists the dissimilarity ratings for the four contrasts of interest. The EX listeners rate all four contrasts as more dissimilar than the IN listeners, but the effect of experience is different for the four contrasts. Table 3 suggests that English language experience increases the sensitivity to [s-z] and also to [θ-f] most, less so to [ð-v], and hardly to [ʃ-ʒ].

Table 3: Mean dissimilarity ratings of four English fricative pairs by inexperienced (top of cell) and experienced (bottom of cell) listeners (range: 0 -8).

English fricatives	Dissimilarity rating
θ-f	1.5 2.5
ð-v	4.6 5.3
s-z	1.5 2.9
ʃ-ʒ	2.2 2.3

The dissimilarity rating yield the following predictions for identification accuracy: EX should outperform IN for [θ], but the low dissimilarity ratings by EX for [f-θ] predict reduced accuracy. The high dissimilarity ratings for [ð-v] by both IN and EX predict fairly accurate identification for both groups, with slightly higher accuracy for EX than IN. The two groups differ greatly in their ratings for [s-z], with low dissimilarity ratings by IN predicting low identification accuracy for [z], and higher ratings predicting more accurate identification by EX. The very similar and quite low dissimilarity ratings for [ʃ-ʒ] by both groups suggest that both IN and EX will have problems identifying [ʒ] correctly.

5. EXPERIMENT 3

Experiment 3 tested the predictions for identification accuracy derived from the perceptual assimilation and the dissimilarity ratings experiments. To allow for direct comparison, Table 4 lists the dissimilarity ratings next to differences in the fit indices for the contrasts of interest. For example, the difference in fit of English [θ-f] to Danish /f/ is (7.2-6.1=) 1.1 for IN, and (6.8-3.3=) 3.5 for EX.

Table 4: Fit index differences (see text) and mean dissimilarity ratings of four English fricative pairs. Top of cell: Inexperienced listeners, bottom of cell: Experienced listeners. (NA because of large percentage of "no fit" responses.)

English fricatives	Fit index difference	Dissimilarity rating
θ-f	1.1 3.5	1.5 2.5
ð-v	NA	4.6 5.3
s-z	1.2 4.2	1.5 2.9
ʃ-ʒ	3.7 1.7	2.2 2.3

Table 4 suggests that the predictions for identification accuracy from two different tasks (perceptual assimilation and dissimilarity rating) are very much the same for three of the four contrasts of interest. Both measures predict problems for IN in identifying English [θ] and [z] because they are perceived to be quite similar to [f] and [s], respectively, and because the difference in fit of English [θ-f] and [s-z] to a native Danish category is very small. The two measures also agree for English [ð-v], which both listener groups perceive to be quite different, and where [ð] is not perceived to fit any native category. The predictions differ, however, for [ʃ-ʒ], where the dissimilarity ratings suggest almost no difference in identification accuracy for IN and EX, but the differences in fit indices should make it easier for IN than EX to identify these fricatives correctly.

These predictions were tested in Experiment 3, which examined the identification of English [f, v, θ, ð, s, z, ʃ, ʒ, w, j, r, l, tʃ, dʒ]. Only the results for the fricatives will be reported.

5.1 Methods

A screen display presented the listeners with 14 orthographic English response alternatives, some of which were given as keywords to disambiguate alternatives, i.e., *f, v, them, think, s, z, sh, genre, ch, jar, w, r, l, y*. After familiarization, participants responded to 84 trials (14 English consonants x 2 talkers x 3 tokens) with an ITI of 0.5 seconds.

5.2 Results and discussion

Table 5 presents the identification matrix for the IN and EX listeners. The EX group was more accurate than IN for almost all fricatives, with the interesting and unexpected exception of [f], a consonant shared by English and Danish, where IN outperformed EX. We suggest that awareness of a difference between [f] and [θ], which is nearly absent in IN and not fully evident in EX, causes uncertainty on the part of EX.

Table 5: Mean percent identification of English /Ca/ tokens by inexperienced (top of cell) and experienced (bottom of cell) listeners. (IN also identified as [v] as /w/ (41.7 %), and [ʒ] as /dʒ/ (27.1%).)

	Response							
	f	v	θ	ð	s	z	ʃ	ʒ
f	95.8 87.5		11.1					
v		57.3 87.5						
θ	72.9 23.6		21.9 68.0					
ð			11.5	71.9 91.7				
s					63.6 90.3	31.3 8.3		
z					35.4 8.3	54.2 90.3		
ʃ							62.5 88.9	25.0 5.6
ʒ							26.1 19.4	36.5 80.6

The results of Experiment 1 and 2 predicted correctly that IN would not identify [θ] correctly, labeling it mostly as /f/. As expected, EX identified [θ] more correctly than IN, but accuracy was still reduced, which is more in line with the low dissimilarity ratings than the large fit index difference for [f-θ]. Both listener groups identify [ð] fairly (IN) or very (EX) accurately, which is well predicted by both the assimilation and dissimilarity ratings for [ð] and [v]. (The low accuracy of IN for [v] is due misidentifications of [w] as /v/, which is not the focus here.) The effect of experience, which was evident in assimilations and dissimilarity ratings for [s] and [z], is reflected in the identification accuracy for these fricatives, which is low for IN and high for EX.

The one surprising result is for [ʃ, ʒ], where the assimilations lead to the counterintuitive expectation that IN would identify these fricatives more accurately than EX, and where the dissimilarity ratings suggested almost no difference between IN and EX. Contrary to the expectations generated by Experiment 1 and 2, EX identified both [ʃ] and [ʒ] highly accurately, whereas IN did not. No explanation can currently be offered for why the effect of experience is reflected only in identification accuracy, but not in the assimilation or dissimilarity ratings of [ʃ] and [ʒ].

6. CONCLUSION

The most important finding of the present study is that both perceptual assimilation and graded discrimination predict identification accuracy nearly equally well. The effect of English language experience was evident in all three experiments, with experienced listeners (EX) exhibiting assimilations and graded discriminations that revealed (in most cases) heightened sensitivity towards English fricatives compared to the inexperienced (IN) listeners. Heightened sensitivity, when evident, was clearly reflected in more accurate identifications. However, the results for [ʃ] and [ʒ] do not fit this overall pattern: Surprisingly, the perceptual assimilations indicated that IN perceived a larger difference between [ʃ] and [ʒ] than EX, which should aid IN in identifying [ʃ] and [ʒ] correctly. That was not the case. Contrary to the expectations generated by the assimilations, EX identified [ʃ] and [ʒ] more correctly than IN. The identification results are somewhat more in line with the graded discriminations, where EX rated [ʃ] and [ʒ] as slightly more different than IN.

Keeping in mind that neither assimilations nor graded discriminations predicted identification of [ʃ] and [ʒ] well, a tentative conclusion from the present study is that dissimilarity ratings are at least as good as perceptual assimilation tasks in predicting identification accuracy. Because previous studies using perceptual assimilation tasks encountered problems when they had to require the use of ambiguous orthographic labels [11], we suggest that future studies consider the potential of dissimilarity ratings (which require no labels) replacing perceptual assimilation tasks as predictors of identification accuracy.

7. REFERENCES

[1] Best, C.T. 1995. A direct realist perspective on cross-language speech perception. In: Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Timonium MD: York Press, 167-200.

[2] Best, C.T., Tyler, M.D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn, O.-S., Munro, M.J. (eds.), *Language Experience in Second language Speech Learning.* In Honor of James Emil Flege. Amsterdam: John Benjamins, 13-34.

[3] Boersma, P., Weenink, D. 2016. Praat, version 6.0.15.

[4] Bohn O.-S., Best, C. T. 2012. Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *J. Phonetics* 40, 109-128

[5] Bohn O.-S., Flege, J.E. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11, 303-328

[6] Flege, J.E. 1995. Second language speech learning: Theory, findings, and problems. In: Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Timonium MD: York Press, 233-277.

[7] Garibaldi, C. L., Bohn, O.-S. 2015. Phonetic similarity predicts ultimate attainment quite well: The case of Danish /i, y, u/ and /d, t/ for native speakers of English and of Spanish. *Proc. 18th ICPhS Glasgow*, 0245

[8] Guion, S. G., Flege, J. E., Akahane-Yamada, R. & Pruitt, J. C. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *J. Acoustical Soc. America* 107(5), 2711-2724.

[9] Horslund, C. S., Ellegaard, A. A., Bohn, O.-S. 2015. Perceptual assimilation and identification of English consonants by native speakers of Danish. *Proc. 18th ICPhS Glasgow*, 0123.

[10] Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E. & Wang, X. 1999. Consonant recordings for speech testing. *J. Acoust. Soc. Am*. 106, L71-L74.

[11] Strange, W., Bohn, O.-S., Trent, S. A. & Nishi, K. 2004. Acoustic and perceptual similarity of North German and American English vowels. *J. Acoust. Soc. Am.*115, 1791-1807