PERCEPTUAL ASSIMILATION AND IDENTIFICATION OF ENGLISH CONSONANTS BY NATIVE SPEAKERS OF DANISH

Camilla S. Horslund[a], Anne A. Ellegaard[b], Ocke-Schwen Bohn[c, d]
Department of English, Aarhus University, Denmark
[c] also at ConAmore, Department of Psychology, Aarhus University, Denmark
[a] engcsh@dac.au.dk, [b] 20115744@post.au.dk, [d] ocke.bohn@dac.au.dk

ABSTRACT

This study explored the relation between the identification of non-native consonants and the perceptual assimilation of non-native consonants to native categories. In Experiment 1, native Danish listeners identified 20 initial English consonants in terms of Danish categories, and then rated the goodness of fit of these matches. The results of this perceptual assimilation experiment were used to predict how accurately Danish listeners would identify English consonants. The predictions were tested in Experiment 2, in which Danish listeners identified the same 20 initial English consonants using English response categories. Our results indicate that perceptual assimilation patterns of non-native consonants predict identification quite well, but they also suggest that current speech learning models need to account for non-native listeners' occasional but systematic response biases.

Keywords: Cross-language identification, nonnative consonants, perceptual assimilation.

1. INTRODUCTION

Models of non-native (L2) speech learning, such as Flege's Speech Learning Model (SLM) [7] and Best's Perceptual Assimilation Model (PAM) [1] and its extension to L2 learning, PAM-L2 [2], base their predictions on the relationship between the phones of the L2 and L1 categories. This relation is often examined in perceptual assimilation experiments in which non-native listeners first identify L2 phones in terms of L1 categories and then rate the goodness of fit of the L2-to-L1 match. This study examines how the perceptual assimilation of the near-complete set of English (EN) initial consonants by native Danish (DK) listeners predicts the identification of EN consonants by DK listeners, thus testing predictions of SLM and PAM as applied to not just individual non-native contrasts as in [4] and [5], but to a range of non-native segments.

SLM views phonetic cross-linguistic similarity as continuous, with L2 phones ranging from *identical* over *similar* to *new*. L2 learners can successfully use L1 categories for identical phones of the L2. Similar phones of the L2, however, differ systematically from their L1 counterpart, but L2 learners are likely to classify them as realizations of the L1 category, a process known as *equivalence classification*, which may lead to incorrect production. New phones are sufficiently different from any L1 category that L2 learners typically notice this difference and eventually establish separate categories for them. One of the basic tenets of SLM is that production problems are perceptually grounded, so despite the fact that SLM was originally a model of L2 speech production, the model has also been applied to L2 perception and will be so for the present study.

PAM states that L2 phones are perceived according to their similarity to and discrepancy from the L1 phone that is closest to them in L1 phonological space. L2 phones may be assimilated to the L1 phonological space as a) a good exemplar of an L1 category, b) an acceptable but not ideal exemplar of an L1 category, c) a notably deviant exemplar of an L1 category, or d) as falling between L1 categories. This results in different assimilation patterns for L2 contrasts to L1 categories: 1) Two-Category (TC) assimilation, when two L2 phones are assimilated to two different L1 categories, 2) Category-Goodness (CG) difference, when two L2 phones are assimilated to the same L1 category, but one is a better exemplar of the L1 category than the other, 3) Single-Category (SC) assimilation, when two L2 phones are assimilated to the same L1 category and are equally good or bad exemplars of that category, and 4) Uncategorised-Categorised (UC) assimilation, when one L2 phone is assimilated to an L1 category while the other L2 phone is not assimilated to any L1 category.

In tests of PAM, perceptual assimilation patterns have been used to predict discrimination of non-native phones, but some studies (e.g. [4]) have also used assimilation results to predict identification for a subset of L1 phones. Whereas discrimination is an appropriate task for naïve listeners, identification requires some knowledge of the L2. Yet, if participants are able to label L2 phones, identification data are more informative than discrimination data, and since identification necessarily involves discrimination, PAM is also applicable to identification studies.

The present study is, to our knowledge, the first to test the predictions of SLM and PAM for the identification of the near-complete set of L2 initial consonants based on the perceptual assimilation of this set. Specifically, we first examined the assimilation of all EN initial consonants (except /m, n, h/) to DK consonants, thus complementing a previous assimilation study of EN vowels to DK [6]. Second, we examined the identification of these consonants by L1 DK listeners. Specific predictions for identification performance were derived from the assimilation results.

A review of the literature on the phonetic realization of initial consonants in EN and DK ([3], [8], [9]) indicates that EN and DK stops both employ short-lag vs. long-lag voicing contrasts and are identical in terms of place of articulation, the only difference being that voiceless DK stops have longer VOTs than their EN counterparts. There is no indication that /s, f, j, m, n, h/ are realized differently in DK and EN. DK has no voiced fricatives - DK /v/ is a labiodental approximant [ʋ], /d/ has a postvocalic approximant allophone, [ð̞][1], and /z/ and /ʒ/ do not exist in DK, and neither does /θ/. EN /ʃ/ and DK /ɕ/ differ in that DK /ɕ/ is more anterior than EN /ʃ/ and lacks lip rounding. EN has the affricates /tʃ, dʒ/, whereas DK has /tj, dj/. DK /l/ is nonvelarized [l], EN /l/ is [ɬ]. EN /r/ is either a "bunched" central dorsal or retroflex approximant [ɻ], whereas DK /r/ is a pharyngeal approximant [ʕ̞][1]. Finally, /w/ is not part of the DK sound system, but may occur in loanwords, where it may be realized as [ʋ] or [w].

Experiment 1 examined how native DK listeners assimilated EN initial consonants to DK initial consonants. The results of Experiment 1 were used to predict how accurately DK listeners would identify EN initial consonants. These predictions were then tested in Experiment 2.

2. EXPERIMENT 1

2.1. Methods

10 L1 DK speakers participated (3 m, 7 f, M_{age} = 22.5 years) as unpaid volunteers. Participants were first year English students at Aarhus University who had spent less than one year in an English-speaking country.

The stimuli were taken from a corpus of EN /Cɑ/ syllables made available by Shannon et al. [11]. We selected 3 tokens each from 2 male L1 American EN speakers in which /C/ was one of the 20 EN consonants [p, b, t, d, k, g, ɻ, tʃ, dʒ, f, v, θ, ð, s, z, ʃ, ʒ, w, j, ɬ].[2] Shannon et al. reported a mean per cent correct identification by L1 EN listeners for the consonants produced by their 5 male talkers of 97.5%.

The TP stimulus presentation software [10] was used to present listeners with 2 randomizations of the (3 tokens x 2 talkers x 20 consonants =) 120 stimuli. Listeners first responded by selecting one of 15 orthographically represented DK response alternatives on a computer screen (i.e., *p, b, t, d, k, g, f, v, s, sj, l, j, r, tj, dj*) plus a "no fit" option. These options correspond unambiguously to DK /p, b, t, d, k, g, f, v, s, ɕ, l, j, r, tj, dj/. Participants then rated the goodness of fit on a scale from 1 (*bad*) to 9 (*good*). Listeners were tested individually in a sound-treated environment using high-quality headphones.

2.2. Results and Predictions

EN plosives were consistently assimilated to their DK counterparts, at rates between 97.5% for [b], 99.2 % for [t, d, k, g], and 100% for [p]. The goodness ratings of these matches were consistently high, ranging from 6.8 for [b] to 7.4 for [p]. Perception of these EN consonants should be unproblematic for DK listeners, because each EN plosive is "identical" (SLM) to its DK counterpart, and because all contrasts should be TC- assimilated (PAM).

Table 1 presents the results for the EN fricatives and affricates. Only EN [f] and [v] were as consistently assimilated to their DK counterparts (and with high goodness ratings) as the plosives. Surprisingly, EN [s], which was uniquely identified with DK /s/, was perceived to be only a moderately good match. Both PAM and SLM would predict that EN [f], [v], and [s] will be perceived highly accurately by DK listeners.

English stimuli	Danish response (and rating)						
	f	v	s	sj	tj	dj	no fit
f	98.3 (7.6)						
v		97.5 (6.8)					
θ	84.2 (7.0)						15.0
ð		48.3 (5.6)					41.7
s			100 (5.3)				
z			84.2 (3.4)				10.8
ʃ				95.8 (5.5)			
ʒ				82.5 (4.3)			
ʧ					95.0 (5.5)		
ʤ					14.2 (4.8)	61.7 (5.4)	

Table 1: Mean percent identification and goodness rating (in parentheses) of EN fricatives and affricates in terms of DK categories.

The predictions are very different for the EN dental fricatives [θ] and [ð]. EN [θ] is perceived to be very similar to DK /f/, and EN [f] and [θ] are either SC assimilated to DK /f/, or CG assimilated with just a very small difference in CG. Both SLM and PAM would predict problems differentiating EN [f] and [θ]. EN [ð], however, is perceived to be quite dissimilar from any DK counterpart ("unassimilated" in terms of PAM), so that both SLM and PAM predict that [ð] will be perceived accurately. The voiced fricatives [z] and [ʒ] are each assimilated to the same DK category as their voiceless counterparts /s/ and /ɕ/. This could suggest perceptual problems, but the difference in category goodness indicates that DK listeners may be able to perceptually differentiate EN [s] – [z] and [ʃ] – [ʒ] to some extent. The moderate goodness ratings for the fit of EN [s] and [ʃ] to their DK counterparts are surprising in the case of [s], but suggest that DK listeners are aware of phonetic differences between EN [ʃ] and DK [ɕ]. The EN affricates are assimilated almost exclusively ([ʧ]) or predominantly ([ʤ]) to contrasting DK categories, which suggests that DK listeners can differentiate them perceptually.

Table 2 presents the results for the English approximants. Comparison with Table 1 reveals that EN [w] and [v] are both assimilated to DK /ʊ/. However, the very large difference in goodness ratings suggests that DK listeners will be able to perceptually differentiate EN [w] and [v]. The EN approximants [ɬ, ɻ, j] are each assimilated to distinct DK categories, which suggest that they will be easily perceptually differentiated. The fairly low goodness ratings for EN [ɻ] and [ɬ] suggest that DK listeners are aware of phonetic differences between EN and DK [ɻ] – [ʕ] and [ɬ] – [l].

English stimuli	Danish response (and rating)				
	v	l	r	j	no fit
w	78.3 (2.9)				
ɬ		80.0 (5.4)			11.7
ɻ			96.7 (4.2)		
j				98.3 (6.3)	

Table 2: Mean percent identification and goodness rating (in parentheses) of EN approximants in terms of DK categories.

3. EXPERIMENT 2

3.1. Methods

10 native speakers of DK (7 f, 3 m, M_{age} = 22.3 years) participated as unpaid volunteers. Participants were English students at Aarhus University who were matched to the participants in Experiment 1 with respect to English language experience.

The TP stimulus presentation software [10] was used to present the same stimuli as in Experiment 1. Listeners were presented with two randomizations and responded by identifying each consonant among 20 orthographically presented alternatives on a computer screen, some of which were given as key words to present unambiguous alternatives, i.e. *P, B, F, V, W, T, D, Think, Them, R, K, G, S, Z, L, Chin, Joke, Ship, Genre, Yes*. Listeners were tested individually in a sound-treated environment using high-quality headphones.

3.2. Results

As expected, the EN plosives were identified highly correctly, with percentages ranging from 96.8 % correct for [b] over 99.2 % correct for [t, k] to 100% correct for [p, d, g].

Table 3 presents the DK listeners' identification accuracy for EN fricatives and affricates. Only the affricates [ʧ, ʤ] were identified highly correctly, which was expected because they were assimilated to contrasting DK categories in Experiment 1. The reduced identifiability of the fricatives is also clearly related to their perceptual assimilation patterns. In Experiment 1, EN [f] and [θ] were both assimilated to DK /f/ with a small difference in goodness of fit. The identification errors reflect this slight difference in category goodness only to some extent because the majority of responses to EN [f] (68.1%) and [θ] (70.8%) were correct. The small goodness-of-fit differences, plus the "no fit" categorization of EN [θ] appears to be sufficient to enable moderately successful identification of EN [f] and [θ].

Correct identification rates for the voiced fricatives [v] (86.7%) and [ð] (85%) are higher than for their voiceless counterparts, which is expected given the assimilation pattern, which indicates that EN [v] is perceived as near-identical to DK /v/, and that EN [ð] is "uncategorized" (PAM) or "new" (SLM). Similarly, the difference in goodness ratings for EN [s, z], which were both assimilated to DK /s/, seems to have enabled the DK listeners to identify the EN sibilants quite accurately. Interestingly, the voiceless [s], which is common to DK and EN, was not identified as correctly as EN [z], which does not exist in DK. A similar result was observed for EN [ʃ, ʒ], which were both assimilated to DK /ɕ/, but with better goodness ratings for [ʃ] than for [ʒ]. Again, the voiced member of the contrast was identified more accurately than the voiceless member. The more accurate perception of the voiced sibilants compared to their voiceless counterparts could be due to hypercorrection; DK listeners' awareness of

the existence of voiced sibilants may induce a hypercorrect response bias.

EN stm	Response									
	f	v	θ	ð	s	z	ʃ	ʒ	ʧ	ʤ
f	68.1		26.7							
v		86.7		9.2						
θ	18.3		70.8							
ð		8.3		85.0						
s					79.2	19.2				
z					9.2	89.2				
ʃ							70.8	24.2		
ʒ							8.3	86.7		
ʧ									98.3	
ʤ										90.8

Table 3: Mean percent identification of EN fricatives and affricates by L1 DK listeners.

Table 4 presents the DK listeners' identification accuracy for EN approximants. As expected, both EN [ɻ] and [j] were identified highly correctly. The lower correct identification rate for EN [j] than [ɻ] could be due to orthographic confusion, as [j] was identified with [dʒ] in 5% of the instances (not shown in Table 4). The relatively low 84.2% correct identification rate of EN [ɬ] is surprising. Not shown in Table 4 are the 9.2% [ð] responses for [ɬ] by the DK listeners, which is difficult to account for. [w] was identified fairly accurately, which was expected given the large difference in goodness ratings for the fit of EN [w] and [v] to DK /v/.

English stimuli	Response			
	w	ɬ	ɻ	j
w	87.5			
ɬ		84.2		
ɻ			100	
j				94.2

Table 4: Mean percent identification of EN approximants by L1 DK listeners.

4. DISCUSSION

The present study examined whether the perceptual assimilation of nonnative consonants, which in previous studies has been used to predict discrimination accuracy, would correctly predict identification accuracy for nonnative consonants. In Experiment 1, L1 DK listeners identified the full set of EN initial consonants (except for [m, n, h]) in terms of DK consonants and rated the goodness of fit of these matches. The results of this perceptual assimilation experiment were used to predict how accurately DK listeners would identify EN consonants using EN labels in Experiment 2.

Experiment 1 revealed that nine EN consonants [p, b, t, d, k, g, ʧ, ɻ, j] were uniquely assimilated to DK consonants with goodness ratings which reflected phonetic differences between DK and EN consonants quite well. Both PAM and SLM would predict high identification accuracy for these EN consonants, which was confirmed in Experiment 2, in which correct identification for these consonants ranged from 94.2% to 100%. Two EN consonants, [ʤ] and [ɬ], were perceptually assimilated less consistently and with only modest goodness ratings to DK counterparts, which however was reflected in reduced identification accuracy only for [ɬ]. The perceptual assimilation patterns for the remaining EN consonants [f, θ, v, ð, s, z, ʃ, ʒ, w] would lead both PAM and SLM to predict reduced identification accuracy because of CG or SC assimilation for [f-θ], [v-ð], [s-z], [ʃ-ʒ], [v-w], or because of the perceived similarity of the members of these contrasts. In general, these predictions were supported, with identification accuracies ranging from 68.1% for [f] to 89.2% for [z]. Surprisingly, identification accuracy was better for EN consonants which in terms of perceptual assimilation and phonetic classification are more different from their DK counterparts [θ, ð, z, ʒ, w] than for those which are perceived to be "identical" (SLM) or "good exemplars" (PAM), i.e., [f, v, s, ʃ]. We suggest that the unexpected relatively high identification accuracy for these new sounds reflects a response bias which is due to some sort of perceptual hypercorrection which works against identifying very similar sounds of the L2 (e.g., [f, s, ʃ]) with their L1 counterparts and instead favors new L2 sounds. The participants of the present study were English students whose education may have made them aware of the novelty of EN [θ, ð, z, ʒ, w] vis-à-vis their L1. It would be interesting to explore whether the apparent response bias shown by these participants would also be observed in DK listeners without such an awareness.

In conclusion, our study showed that perceptual assimilation of a (near-complete) set of nonnative consonants predicts their interlingual identification quite well. The apparent bias skewing responses towards the most dissimilar members of (mostly) fricative contrast was not anticipated and should be explored in future studies.

1 The IPA has no separate symbol for this; approximant is indicated by using a lowering diacritic below the homorganic voiced fricative symbol.

2 We did not include [m, n, h] among the stimuli because piloting revealed that EN [m, n, h] were always assimilated to DK /m, n, h/, respectively, and that DK listeners identified these consonants 100% correctly.

5. REFERENES

[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] Bohn, O.-S. 2013. Acoustic characteristics of Danish infant directed speech. *Proc. Meetings on Acoustics* 19, 060055.

[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] Bohn, O.-S., Steinlen, A. K. 2003. Consonantal Context affects Cross-Language Perception of Vowels. *Proc. 15th ICPhS Barcelona*, 2289-2292.

[7] 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.

[8] Garibaldi, C., Bohn, O.-S. Phonetic similarity predicts ultimate attainment in adults quite well: The case of Danish /y/ and /d/, /t/ for native speakers of English and of Spanish. Submitted.

[9] Grønnum, N. 1998. Danish. *J. Intern. Phonetic Association* 28, 99-105.

[10] Rato, A., Rauber, A., Kluge, D., Santo, G. 2013. Designing Audio, Visual, and Audiovisual Perceptual Experiments with TP Software. *Poster presentation at NEW SOUNDS 2013, May 17-19*, Concordia University, Montreal, Canada.

[11] 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.