

Consonantal Context affects Cross-Language Perception of Vowels

Ocke-Schwen Bohn¹, Anja K. Steinlen²

¹ Aarhus University, Denmark

² Kiel University, Germany

E-mail: engosb@hum.au.dk

ABSTRACT

This study examined whether and how cross-language perception of vowels is affected by the consonantal context in which vowels occur. Native Danish (DK) listeners identified the 11 monophthongs of Southern British English (SBE), produced in three consonantal contexts (/hVt/, /dVt/, /gVk/), using DK response categories, and the listeners provided goodness ratings of their interlingual identifications. Responses indicate that perceptual assimilation of some vowels is strongly affected by consonantal context. Results are interpreted by referring to the results of an acoustic study [1] which showed that the location of DK vowels in the F1/F2 space is minimally affected by flanking consonants, whereas the center frequencies of SBE vowels are strongly affected by consonantal context. Apparently, DK listeners transfer their native expectations about (lack of) vowel coarticulation to SBE. Our results suggest that patterns of perceptual assimilation (and predictions of learning difficulty) of nonnative vowels depend upon the phonetic contexts of vowels.

1. INTRODUCTION

Current models of cross-language and second language speech perception [2, 3] assume that ease or difficulty in perceiving non-native sounds can be predicted on the basis of how non-native sounds map on to native categories. With a few recent exceptions [e.g., 4, 5, 6], most cross-language and second language studies have examined interlingual identification and perceptual assimilation of nonnative to native sounds using citation-form stimuli produced in only one phonetic context (e.g., [7]). However, two broad experimental findings mentioned by Beddor et al. [8] question the assumption that the relation of native to nonnative sounds can be generalized from studies which present nonnative sounds in just one phonetic context. These findings are (a), with respect to speech production, that languages differ systematically in their patterns of coarticulation (as reviewed in [9]) and (b), with respect to speech perception, that listeners are sensitive to coarticulatory variability (as reviewed in [10]).

These findings make it seem likely that listeners have language-specific expectations about coarticulation which are well-attuned to their native language (L1), preserving

perceptual constancy in the face of L1-specific coarticulatory variability. However, these L1-specific expectations may not serve the listener well when encountering nonnative speech sounds which exhibit different coarticulatory patterns. For example, Steinlen [1] reported that SBE and DK differ considerably in the extent to which the place of articulation of flanking consonants in consonant-vowel-consonant (CVC) syllables affects the acoustic identity of vowels, i.e., their location in the F1/F2 space. Specifically, the location of DK vowels was minimally affected by whether they were produced in the “neutral” /hVt/ context or with bilabial, alveolar, or velar flanking stop consonants. The center frequencies of SBE vowels, however, differed systematically (and, for some vowels, considerably) across different consonant contexts.

Steinlen’s findings lead to the expectation that perceptual assimilation of nonnative vowels (and hence learning difficulty) may not be uniform across phonetic context. To our knowledge, this assumption has previously been tested in only two studies [5, 6]. To further test this assumption with two languages which differ greatly in how coarticulation affects acoustic vowel identity, the present study examined how native DK listeners categorized the monophthongs of SBE using DK response categories, and how they judged the goodness of fit of the SBE vowel to the DK category, when the SBE vowels had been produced by multiple speakers in different consonantal contexts.

The SBE vowel inventory contains 11 stressed monophthongs: /i:, ɪ, ɛ, æ, u:, ʊ, ɔ:, ɒ, ɑ:, ɜ:, ʌ/. As indicated by the length diacritic, neighboring vowels differ both in vowel quality and in duration. DK has 20 stressed monophthongs: /i:, i, e:, e, ɛ:, ɛ, æ:, æ, y:, y, ø:, ø, œ: œ, u:, u, o:, ɔ, ɔ:, ɔ/. The symbols used for DK vowels correctly indicate that DK vowels can be grouped in long/short pairs which do not differ much in vowel quality [1, 11]. It should be noted, however, that the use of the same symbol for DK and SBE vowels can be misleading because almost all DK vowels are located in the upper half of the acoustic vowel space. For example, Steinlen’s acoustic measurements of SBE /ɪ/ and DK /ɛ/ revealed that these two vowels did not differ significantly in their formant frequencies if produced in /hVt/ syllables. - Some DK vowels have allophones that are phonetically quite distinct, in particular, postvocalic /r/ is realized with lowered and lengthened variants for the preceding vowel (Danish is non-rhotic), and labial and velar environments cause /æ/ to be realized as [ɑ].

2. METHODS

2.1 Subjects

30 native speakers of DK (21 f, 9 m, mean age: 24.3 years) participated as unpaid volunteers. The subjects were first- and second-year students at the English Department of Aarhus University who had minimal English-language experience (less than 6 months of residence in an English speaking country), were speakers of the East Jutland/Aarhus dialect of DK, and who had SBE as their target variety (all according to self-report elicited in a language background questionnaire).

2.2 Stimuli

The stimuli represent a subset of the tokens used by Steinlen [1] to examine language-specific patterns of vowel coarticulation with flanking consonants in CVC syllables. In the Steinlen study, 10 native speakers each of SBE, DK, and North German produced native monophthongs in a variety of contexts differing in place of articulation of the flanking consonants. For the present study, the productions by 4 male native speakers of the 11 monophthongs of SBE (/i:, ɪ, ɛ, æ, u:, ʊ, ɔ:, ɒ, ɑ:, ɜ:, ʌ/) in /hVt/, /dVt/, and /gVk/ syllables were selected. The stimulus set consisted of 3 repetitions by each speaker of each CVC syllable, i.e., of 3 x 4 x 11 (= 132) stimulus syllables for each of the 3 contexts.

2.3 Procedure

Three audio cassette tapes, one for each context, were generated with a randomized set of either /hVt/, /dVt/, or /gVk/ syllables. Each stimulus syllable was recorded twice, once for interlingual identification, and, immediately following, for category goodness rating. Stimuli were presented with an ISI of 3.0 sec, and there was a 5.0 sec break after each block of 12 presentations. The 30 subjects were randomly assigned to 3 groups which listened to the SBE vowels in either the /hVt/, /dVt/, or /gVk/ condition. Each subject listened individually to the stimuli which were presented over Sennheiser HD450 headphones from a Marantz CP430 cassette player in a quiet environment at a self-selected comfortable hearing level. Each subject received a response booklet with 12 pages which each contained 12 lines. Each line listed the response alternatives in Danish orthography and the numbers 1 to 5 to indicate the goodness of fit for each interlingual identification. The first page of the response booklet was used for familiarization. A separate stimulus tape was used for familiarization, and the responses were not used in data analysis. The 11 response alternatives in each line of the response booklet had been determined from the results of open-set responses in pilot experiments.

For the /hVt/ context, DK response alternatives were *hidt, hedt, hædt, hadt, hødt, hådt/hodt, hudt, hardt, hørdt, hårdt* (containing the vowels [i, e, ɛ, æ, ø, ɔ, u, ɑ:, œ:, ɒ:]). For the /dVt/ context, DK response alternatives were *dit, dett, dætt, dat, døt, dåt/dot, dut, dart, dørt, dort* (containing the vowels [i, e, ɛ, æ, ø, ɔ, u, ɑ:, œ:, ɒ:]). For the /gVk/ context, DK response alternatives were *gik, gek, gæk, gak, gøk, gåk/gok, guk, gark, gørk, gork* (containing the vowels [i, e, ɛ, a, ø, ɔ, u, ɑ:, œ:, ɒ:]). Note that the phonetic manifestation of DK /æ/ in the velar context is [a]. Each line also listed the

response category *andet* (Engl. “other”) which the listeners were encouraged to use if none of the other categories provided a match. In that case, they were asked to write their response. The end of each line listed the numbers 5, 4, 3, 2, 1 for goodness ratings, with the endpoints 5 labeled as *passer godt* (Engl. “fits well”) and 1 labeled as *passer dårligt* (Engl. “fits badly”).

Before familiarization, subjects were instructed to read the first line of response alternatives out loud. The orthographic rendition of the DK response alternatives, most of which were not real DK words, was phonetically unambiguous. Subjects were encouraged to use the “other” category (if necessary), and to use the entire five-point scale for their goodness ratings.

The fact that the short vowels of SBE can only occur in closed syllables, whereas the long vowels of DK can only occur in open syllables could be seen to create a problem insofar as the DK closed-syllable response alternatives deprived listeners of the choice to respond with DK long vowels (except for the responses with postvocalic /r/, i.e., [ɑ:, œ:, ɒ:]). Because the full set of SBE monophthongs can only be produced in closed syllables, this potential problem could not be resolved. Note, however, that DK vowel *qualities* are much the same for long/short pairs.

3. RESULTS

The tables below present the assimilation patterns for SBE vowels produced in the /hVt/ context (Table 1), the /dVt/ context (Table 2), and the /gVk/ context (Table 3). Each table gives the stimulus SBE vowel and the most frequent and second most frequent DK response category (if the second most frequent response category was chosen in more than 10% of the instances). Each table also lists the frequencies of selection of responses (converted to percentages of total opportunities, in parentheses), and the median goodness ratings assigned to the SBE-DK matches. The percentages reflect the consistency with which listeners identified each SBE vowel with a particular DK response category, and the goodness ratings indicate the perceived goodness of fit of the SBE vowel to the selected DK category. Rows that are marked in bold face indicate an inconsistent choice of DK response categories across consonantal contexts.

The tables show that the DK listeners varied considerably in both the choice of DK response categories for individual SBE vowels and in the consistency with which they identified SBE vowels with DK categories. Overall, the responses of the DK listeners seem to fall into three classes: Those that were quite or highly consistent (for SBE /i:, u:, æ, ɒ, ɑ:/, with response consistency ranging from 87% to 100%), those that were less consistent but not much affected by phonetic context (SBE /ɔ:, ɜ:/, with response consistency ranging from 68% to 86%), and those that were quite inconsistent and considerably affected by phonetic context (SBE /ɪ, ɛ, ʊ, ʌ/, with response consistency ranging from 37% to 81%).

DK response category		
SBE vowel	Most frequent	Second most frequent
i:	i (99) 3	
ɪ	e (64) 4	i (28) 3
ɛ	æ (55) 4	ɛ (25) 4
æ	æ (100) 4	
u:	u (100) 3	
ʊ	u (67) 3	o (20) 3
ɔ:	ɒ: (80) 4	ɔ (19) 3
ɒ	ɔ (96) 3	
ɑ:	ɑ: (89) 4	
ɜ:	œ: (72) 2	ø (23) 3
ʌ	ɔ (37) 3	æ (20) 4

Table 1: Perceptual assimilation patterns for SBE vowels produced in an /hVt/ context. The cells list the most frequent and the second most frequent DK response category along with the response consistency (in parenthesis) and the median goodness ratings (1 = poor, 5 = good). See text for details.

DK response category		
SBE vowel	Most frequent	Second most frequent
I:	i (97) 3	
ɪ	e (57) 3	i (23) 2
ɛ	æ (53) 4	ɛ (40) 3
æ	æ (98) 4	
u:	u (95) 2	
ʊ	ø (41) 3	ʊ (33) 2
ɔ:	ɒ: (68) 3	ɔ (21) 3
ɒ	ɔ (98) 4	
ɑ:	ɑ (93) 4	
ɜ:	œ: (80) 2	ø (12) 3
ʌ	ɔ (42) 3	æ (20) 3

Table 2: Perceptual assimilation patterns for SBE vowels produced in a /dVt/ context. See text for details.

DK response category		
SBE vowel	Most frequent	Second most frequent
i:	i (97) 4	
ɪ	i (52) 4	e (38) 4
ɛ	ɛ (76) 4	e (14) 3
æ	a (96) 3	
u:	u (99) 4	
ʊ	u (81) 4	
ɔ:	ɒ: (86) 4	ɔ (11) 1
ɒ	ɔ (87) 4	
ɑ:	ɑ (88) 4	
ɜ:	œ: (78) 3	ø (17) 1
ʌ	ø (40) 3	ɔ (18) 3

Table 3: Perceptual assimilation patterns for SBE vowels produced in a /gVk/ context. See text for details.

Five of the SBE monophthongs were identified quite or highly consistently with DK counterparts: /i:/ (97%-99% consistency), /u:/ (95%-100% consistency) /æ/ (96%-100% consistency), /ɒ/ (87%-98% consistency), and /ɑ:/ (88-93% consistency). (The perceptual assimilation of SBE /æ/ is a special case because DK /æ/ is [a] in the velar /gVk/ context. The relatively low median goodness rating for SBE /æ/ -> DK [a] in /gVk/ suggests that the DK listeners are aware that this match is not as good as for SBE /æ/ -> DK [æ] in the /hVt/ and /dVt/ contexts.) Interestingly, the most consistent perceptual assimilations were observed for the four SBE vowels which define the SBE vowel space. Note, however, that the goodness ratings for the corner vowels /i:/ and /u:/ indicate that the DK listeners were sensitive to differences between these SBE and DK corner vowels in different phonetic contexts: The median goodness ratings for SBE /i:/ -> DK [i] ranged from a fair 4 in the /gVk/ context to a mediocre 3 in the /hVt/ and /dVt/ context, and the median goodness ratings for SBE /u:/ -> DK [u] ranged from a fair 4 (/gVk/) over 3 (/hVt/) to a quite poor 2 (/dVt/). The poor match that DK [u] provided for SBE /u:/ was most likely due to the fact that SBE /u:/ is a fronted [ɥ] in the alveolar context (see [1]); the 5% “other” responses from the DK listeners for SBE /du:t/ were all DK [y].

Two SBE vowels, /ɔ:/ and /ɜ:/, were identified less consistently with DK vowels than the corner vowels, but their perceptual assimilation was not much affected by phonetic context. The perceptual assimilation of SBE /ɔ:/ to DK [ɒ:] as the modal response category and to DK [ɔ] as the second most frequent response suggests that the DK listeners were more influenced by the duration of SBE /ɔ:/ than by its spectral properties because this SBE vowel is

spectrally more similar to the short DK [ɔ] vowel than to long DK [o:]. The low goodness ratings (1-3) assigned to the DK counterparts [œ:] and [ø] for SBE /ɜ:/ indicates a sensitivity to the considerable spectral differences between the unrounded SBE and the rounded DK vowels.

Perceptual assimilation of the SBE vowels /ɪ, ɛ, ʊ, ʌ/ was strongly affected by consonantal context. These vowels were not each uniquely identified with a DK response category. Rather, the most frequent DK responses differed either in the velar or in the alveolar context from the responses in the other two contexts. Comparison of the perceptual results obtained here to the acoustic results of Steinlen's [1] study, which compared the center frequencies of SBE and DK vowels in various consonantal contexts, suggests that the DK listeners performed the perceptual assimilations of the SBE vowels /ɪ, ɛ, ʊ, ʌ/ on the basis of native expectations about minimal coarticulatory effects of flanking consonants on vowels in DK CVC syllables.

Specifically, the most frequent response to SBE /ɪ/ was DK [e] in the /hVt/ and /dVt/ contexts, but DK [i] in the /gVk/ context, probably because the /gVk/ context induces raised realizations of /ɪ/ in SBE, whereas (the center frequencies of) DK vowels are not affected by a velar context, see [1]). Likewise, and probably for the same reason, the most frequent responses for SBE /ɛ/ was DK [ɛ] when presented in /gVk/ syllables, but DK [æ] when presented in /hVt/ and /dVt/ syllables. Similarly, the most frequent response to SBE /ʊ/ was DK [u] in the /hVt/ and /gVk/ contexts, but DK [ø] in the /dVt/ context, probably because SBE /ʊ/ was considerably fronted in the alveolar context. – These results, plus the fact that SBE /ʌ/ was the least consistently identified vowel (e.g., 5 DK response categories with > 10% responses in the /hVt/ condition) suggests that inconsistency in perceptual assimilation may entail susceptibility to context effects.

4. CONCLUSIONS

This study provided further evidence that the perceptual assimilation of vowels cannot be predicted from comparisons of phonological inventories or from studies which examine cross-language perception in just one phonetic context. The DK listeners perceptually assimilated SBE vowels to different spectral categories depending on the context in which vowels occurred, and/or the context affected the perceived goodness of fit of the interlingual identifications. Our results can be used to predict difficulty in cross-language perception and in the acquisition of SBE vowels by L1 DK speakers, and we predict that this difficulty will vary considerably as a function of the context in which vowels are presented.

REFERENCES

- [1] Anja K. Steinlen, A cross-linguistic comparison of the effects of consonantal contexts on vowels produced by native and non-native speakers. Unpublished doctoral dissertation, Aarhus University, 2002.

- [2] Catherine T. Best, "A Direct-Realist View of Cross-Language Speech Perception", In Winifred Strange (ed.), *Speech Perception and Linguistic Experience*. Baltimore: York Press, 1995, pp. 171-204.
- [3] James E. Flege, "Second Language Speech Learning: Theory, Findings, and Problems", In Winifred Strange (ed.), *Speech Perception and Linguistic Experience*. Baltimore: York Press, 1995, pp. 233-277.
- [4] Winifred Strange, Reiko Akahane-Yamada, Rieko Kubo, Sonja A. Trent, Kanae Nishi, and James J. Jenkins, "Perceptual assimilation of American English vowels by Japanese listeners", *Journal of Phonetics*, **26**, pp. 311-344, 1998.
- [5] Winifred Strange, Reiko Akahane-Yamada, Rieko Kubo, Sonja A. Trent and Kanae Nishi, "Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners", *Journal of the Acoustical Society of America*, **109**, pp. 1691-1704, 2001.
- [6] Winifred Strange, Ocke-Schwen Bohn, Sonja A. Trent and Kanae Nishi, "Contextual variation in the acoustic and perceptual similarity of North German and American English vowels", submitted for publication, 2002
- [7] Ocke-Schwen Bohn and James Emil Flege, "The production of new and similar vowels by adult German learners of English", *Studies in Second Language Acquisition* **14**, pp.131-158, 1992.
- [8] Patrice S. Beddor, James D. Harnsberger and Stephanie Lindemann, "Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates", *Journal of Phonetics*, **30**, pp. 591-627, 2002.
- [9] Sharon Manuel, "Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts", In William J. Hardcastle and Nigel Hewlett (eds.), *Coarticulation: Theory, data, and techniques*. Cambridge, UK: Cambridge University Press, 2000, pp. 179-198.
- [10] Burton S. Rosner and J. Pickering, "Vowel perception and production. Oxford, UK: Oxford University Press, 1994.
- [11] Eli Fischer-Jørgensen, "Formant frequencies of long and short Danish vowels", *Annual Reports of the Institute of Phonetics, University of Copenhagen*, **7**, pp. 21-53, 1972.

ACKNOWLEDGMENT

Supported by grants from Aarhus Universitets Forskningsfond and Statens Humanistiske Forskningsråd.