Summary

Second-language speech perception and production in adult learners before and after short-term immersion

Ph.D. dissertation by
Anders Damgren Højen
University of Aarhus
2003

Introduction

The experimental research reported in this thesis examined English speech learning in Danish adult learners of English in their initial phase of extensive exposure to native English conversational speech. The subjects' perception and production of English and Danish were tested before and after short-time immersion in Southern England.

Several studies have reported that added second-language (L2) experience results in a more native-like L2 speech performance in adult L2 learners. The amount of experience has often been quantified in terms of the length of residence (LOR) in an L2 speaking community. While some studies reported an effect of LOR on L2 performance (e.g., Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997; Flege & Hillenbrand, 1984; Yamada, 1995) other studies reported no effect of LOR (e.g., Flege, 1988, 1993; Flege, Munro, & Skelton, 1992).

As an explanation for the lack of a significant difference between experienced and inexperienced learners in Flege's (1988) foreign accent study, Flege & Fletcher (1992) suggested that the "inexperienced" learners in were too experienced, i.e., they had come too close to their ultimate level of pronunciation proficiency. This, in turn, led the authors to suggest that most improvements in L2 pronunciation occur during the first year of exposure to the L2. Piske et al. (2001), in a literature review, supported the conclusion drawn by Flege (1988) and Flege & Fletcher (1992) that L2 learning proceeds at a faster rate in the initial learning period.

In order to test the hypothesis that one year of extensive L2 input would result in substantial phonetic L2 learning, and in order to examine learning in a longitudinal design, the present study examined the production and perception of English speech

by adult native speakers of Danish before and after a stay of 3-11 months in Southern England.

Methods

The study had a longitudinal design. Three subject groups were formed. The Experience Group consisted of 14 native Danish female au pairs or students (mean age = 21 years), who went to Southern England for an average of 7.1 months (range = 3-11). This group was tested before (Time 1) and after (Time 2) their stay abroad. The No-experience Group consisted of 11 native Danish females (mean age = 20 years) who stayed in Denmark and served as a control group. This group was also tested twice. The Native English Group consisted of 6 females (mean age = 20 years) who were tested once and provided English baseline data.

The subjects participated in production tests and perception tests. In experiment 1, an English sentence produced at Time 1 vs. Time 2 was rated for global foreign accent by ten native English judges. In Experiment 2, the production and perception of English [s] and [ʃ] were examined through listener judgments, acoustic analyses, and perception tests. In Experiment 3, the production and perception of English [ɑ] and [ʌ] were examined in the same way.

Results and Discussion

In Experiment 1, as indicated in Figure 1, the subjects in the Experience Group received significantly better foreign accent scores at Time 2 than Time 1 [$t = 2.790$, df. 13, $p < 0.007$] (one-tailed prediction). However, their mean foreign accent score was not within the range of scores assigned to the Native English Group.

In agreement with Hypothesis 1, there was a significant correlation between LOR and increase in foreign accent score from Time 1 to Time 2 [$r(12) = 0.606$, $p<0.022$]. However, there was an even stronger correlation between score increase and a measure of "total input", (LOR weighted by amount of English use) [$r(12) = 0.811$, $p<0.000$], see Figure 2. This finding suggests that LOR may not be a sufficiently accurate index for linguistic experience, but should be weighted by amount of L2 use.

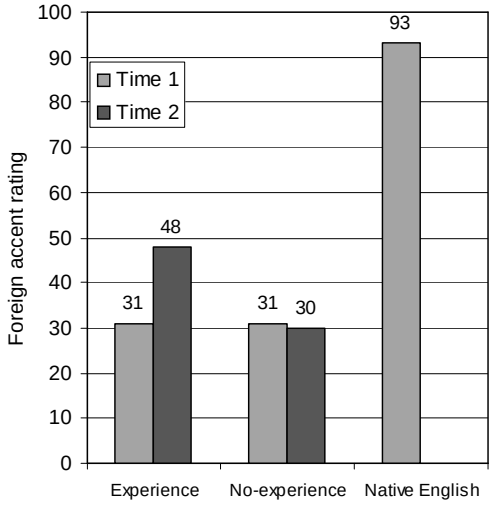

Figure 1 The mean foreign accent scores given to each subject Group. A score of 100 means "no foreign accent; a score of 1 means "strong foreign accent".

The significant effect of L2 experience after just 7.1 months suggests that much learning takes place during the first year of extensive exposure to conversational L2 speech as suggested by Flege & Fletcher (1992). However it is not possible to conclude from the present study whether *most* learning takes place during the first year as suggested by Flege & Fletcher (1992). Future longitudinal studies should extend over several years in order to determine when an asymptote in the learning curve occurs.

The fact that the Experience Group received significantly higher scores after a relatively short period of immersion in an English-language environment and the fact that one subject achieved native-like pronunciation ratings do not appear to be consistent with the Critical Period Hypothesis. One of the proponents of the Critical Period Hypothesis is Long (1990), who asserted that

> The ability to attain native-like phonological abilities in a SL [second language] begins to decline by age 6 in many individuals and to be beyond anyone beginning later than age 12, no matter how motivated they might be or how much opportunity they might have.

Select speech sounds were also examined through perception and production tests. In Experiment 2A and 2B, it was expected that English [ʃ] would be phonetically different from Danish [ɕ] but still be perceived as equivalent to Danish [ɕ]. It was hypothesized that, in terms of Flege's (1995) Speech Learning Model (SLM), English [ʃ] would be perceived as so similar to Danish [ɕ] that the Experience Group subjects would (begin to) develop a diaphone category for English [ʃ] and Danish [ɕ]. Due to English-language input during immersion, the diaphone should

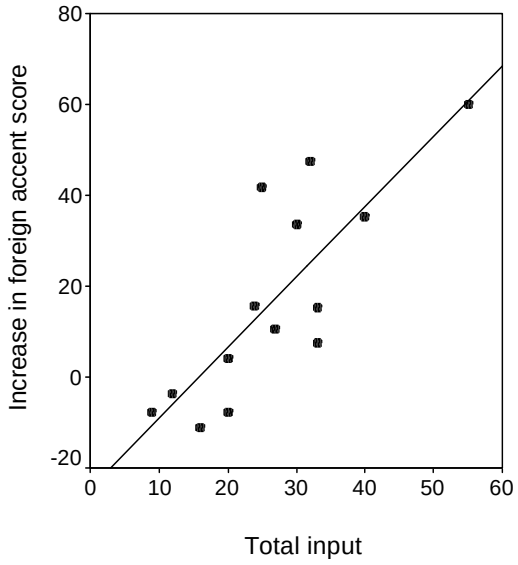

Figure 2. The correlation between total input (LOR x English use) and foreign accent score increase from Time 1 to Time 2.

develop to reflect the characteristics of both sounds.

Thus, the representation of the diaphone in the long-term memory of the immersed learners should become more [ʃ]-like at Time 2. This change was expected to be revealed in a displaced perceptual category boundary between [s] and [ʃ] in both Danish and English language sets and both Danish and English words should be produced with more [ʃ]-like characteristics.

However, it turned out to be difficult to interpret both the perception and the production results in terms of the SLM framework due to unexpected results. The perception experiment supported the hypothesis that the perceptual category boundary between [s] and [ʃ] would move in the direction of English [ʃ] at Time 2 in the English as well as the Danish language set. However, a trend in the same direction for a subset of the No-experience Group indicated that the perceptual change could have had other causes than added English-language experience.

Another unexpected finding was that both the Experience Group and the No-experience Group produced English [ʃ] authentically already at Time 1. Even less expected was the finding that Danish [ɕ] did not differ acoustically from native English [ʃ]. This finding could be either due to a recent sound change in Danish [ɕ] towards [ʃ], or, more likely, due to imitation. Immediately prior to producing Danish [ɕ], the L1 Danish subjects had perceived a Danish *sok-sjok* continuum in which they heard Danish [ɕ] spoken with very low-frequency spectral characteristics as endpoint stimuli (i.e., similar to [ʃ]). This sound is likely to sound like a hyper-articulated [ɕ] to L1 Danes, and since subjects probably want to perform "well" in experiments, the L1

Danes might have hyper-articulated Danish [ɕ] according the examples of [ʃ] which they had heard in the perception tests.

More studies are needed to explore the two outlined explanations (sound change or imitation) for the similarity of Danish [ɕ] and English [ʃ] in the present study. However, it can be concluded that Danish learners of English probably have no problems producing English [ʃ] authentically.

Experiment 3A and 3B tested the hypothesis that L1 Danish learners of English would not be able to perceive or produce a difference between English [ɑ] and [ʌ], but that exposure to authentic conversational English speech would enable L1 Danes to

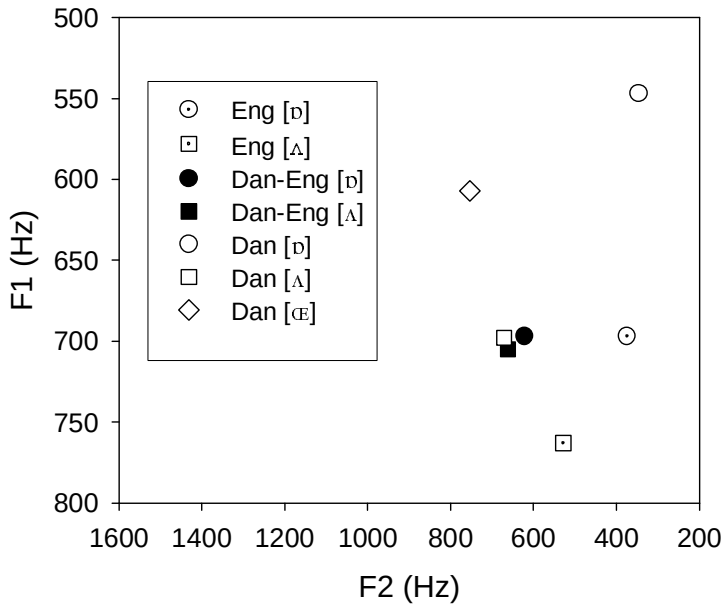

Figure 3. Mean spectral values of the English vowels [ɒ] and [ʌ] and the Danish vowels [ɒ], [ʌ] and [œ] plotted in the first formant by second formant minus first formant vowel space. All vowels were spoken at Time 1.

Eng: English vowels spoken by the Native English Group.
Dan-Eng: Danish accented English vowels spoken by the Experience Group.
Dan: Danish vowels spoken by the Experience Group.

do so.

The results showed that the Experience Group had not learned to produce a difference between English [ɑ] and [ʌ] at Time 2. On the other hand, both L1 Danish

groups appeared to perceive the vowel contrast categorically already at Time 1; in the perception of a resynthesised [ɑ]-[ʌ] continuum, the Danes showed an identification cross-over and a discrimination peak, though not as categorical as those of the Native English Group.

The fact that the L1 Danes could not produce a difference between English [ɑ] and [ʌ] suggested that they perceived both sounds as equivalent to one Danish vowel category. This was further supported by the fact the English [ɑ] and [ʌ] produced by the Experience Group were acoustically very similar to their Danish [ʌ] (see Figure 3). This suggests that the English [ɑ]-[ʌ] contrast in the framework of the Perceptual Assimilation Model (PAM, Best, 1995) is either a Single-Category or, more likely, a Category-Goodness assimilation type for L1 Danes. The L1 Danish subjects' relatively successful discrimination and categorical identification of English [ɑ] and [ʌ] is consistent with the contrast being a CG type, for which PAM predicts "moderate" to "very good" discrimination.

However, if assimilation type is determined on the basis of obtained discrimination, the definitions become circular. On the assumption that English [ɑ]-[ʌ] was a CG contrast, the finding that the Experience Group developed a more categorical perception of the contrast after L2 experience supports a finding reported by Guion et al. (2000), who examined the discrimination of different PAM contrasts by groups of L1 Japanese learners of English who varied in English-language experience. Guion et al. reported that discrimination of a CG contrast improved with L2 experience, whereas discrimination of a both uncategorized (UU) type of contrast did not differ across varying levels of experience.

The finding that the L1 Danish speakers were relatively successful at identifying and discriminating English [ɑ] and [ʌ] is consistent with the Native Language Magnet (NLM) model (Kuhl & Iverson, 1995). According to the NLM model, two non-prototypical members of a category are discriminated more easily than a prototype vs. a non-prototype, because the two non-prototypes evade the perceptual magnet effect. Further support for the magnet effect would be provided if a study showed less

categorical perception in Danish learners' identification and discrimination of a continuum ranging from the Danish [ʌ] prototype to English [ɑ] or to English [ʌ].

Experiment 3A and 3B supported Hypothesis 4 that English language experience would have a stronger effect on the perception than the production of L2 speech sounds. The Experience Group perceived the English [ɑ]-[ʌ] contrast more categorically at Time 2 than at Time 1 (though not as categorically as the Native English Group did). On the other hand, the Experience Group had not learned to produce a difference between [ɑ] and [ʌ] at Time 2. These results suggest that perception leads production in the learning of L2 sounds.

Conclusions and future research

The present study yielded results and raised new questions that have implications for future research in terms of both methodology and theory. In Experiment 1, it was found that a mean LOR of 7.1 years was enough to improve the perceived foreign accent in L1 Danish speakers of English. This suggests that future studies on the effect of LOR should take into account the fact that the learning curve seems to be steep in the first year of exposure to L2 speech (given sufficient L2 use). Flege & Fletcher (1992) suggested that most improvement in L2 speech learning takes place during the first year of extensive exposure. Future longitudinal research should extend over several years to determine when an asymptote in the L2 speech learning is reached.

As predicted, LOR correlated significantly with increase in foreign accent score. However, the composite measure *Total input* (LOR weighted by amount of English use) was even more successful at predicting the increase in foreign accent score. This suggests that LOR alone is not an accurate index of L2 experience, supporting proposals by Flege and Liu (2001).

The finding in Experiment 2B that the L1 Danish learners unexpectedly produced English [ʃ] authentically was assumed to be due to imitation of stimuli in a preceding perception task. If this is correct, English [ʃ] might be easy for native Danes to acquire.

The Danish learners of English could not produce a difference between [ɑ] and [ʌ] suggesting that they did not have separate phonetic categories for the two vowels. But still, they were relatively successful at identifying and discriminating the two English vowels. This is unexpected on a traditional categorical perception view if the two English vowels are not perceived as categorically different. However, the result might be explained in terms of the NLM model (Kuhl & Iverson, 1995) which suggests that non-prototypical instances are of a category are more discriminable than a prototype vs. a non-prototypes. This possibility could be further examined by testing native Danish speakers' ability to discriminate continua ranging from Danish [ʌ] to English [ɑ], or from Danish [ʌ] to English [ʌ]. According to the NLM model, such continua should not be perceived categorically by L1 Danes if one member of the pair is similar to Danish [ʌ].

Experiment 3B showed that the Experience Group had not learned to produce a difference between English [ɑ] and [ʌ] at Time 2. The difficulty that native Danish speakers have with this vowel contrast could arise in part from the fact that the lower central and back part of the Danish vowel space is occupied by conditioned allophones. Future research in how L1 Danes perceive and produce English low back vowels in different phonetic contexts that condition different allophones could provide further insight into the importance of phonotactical constraints in the perception of L2 sounds. Specifically, it would be interesting to examine how L1 Danes perceive and produce English [ɑ] and [ʌ] in different phonetic contexts which resemble the contexts in which the contextually conditioned Danish allophones [ʊ], [ʌ], [ɔ], and [ɶ] (an allophone of [œ]) may occur.

In summary, the findings suggest that the phonetic system in adults is still malleable in young adults and that perception leads production in L2 speech acquisition. The finding that global foreign accent ratings improved but that no improvement was found in the select speech sounds might suggest that important improvements happened in the prosodic dimension.

References

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* (pp. 171-204). Timonium, MD: York Press.

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.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*(1), 70-79.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America, 93*(3), 1589-1608.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on the degree of perceived foreign accent. *Journal of the Acoustical Society of America, 91*(1), 370-389.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America, 76*(3), 708-721.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults acquisition of a second language. *Studies in Second Language Acquisition, 23*(4), 527-552.

Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America, 92*(1), 128-143.

Guion, S. G., Flege, J. E., Yamada, R. A., & Pruitt, J. C. (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*(5), 2711-2724.

Kuhl, P. K., & Iverson, P. (1995). *Linguistic experience and the "Perceptual Magnet Effect"* (1st ed.). Timonium, MD: York Press.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251-285.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics, 29*(2), 191-215.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds. Perception of American English /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 305-320). Timonium, MD: York Press.