Schouten, M.
1975 *Native-language Interference in the Perception of Second-Language Vowels*. Utrecht: Drukkerij Elinkwijk B.V.
Smith, B.
1977 "Temporal aspects of English speech production: a developmental perspective", *Journal of Phonetics* 6, 37–69.
Snow, C.
1988 "Relevance of the notion of a critical period to language acquisition", in M. Bornstein (ed.), *Sensitive Periods in Development: Interdisciplinary Perspectives*. Hillsdale N.J., Lawrence Erlbaum, 183–210.
Snow, C. – M. Hoefnagel-Höhle
1977 "Age differences in the pronunciation of foreign sounds", *Language and Speech* 20:357–365.
Snow, C. – M. Hoefnagel-Hohle
1982 "Age differences in the pronunciation of foreign sounds", in R. Scarcella and M. Long (eds.), *Child-Adult Differences in Second Language Acquisition*. Rowley MA: Newbury House, 84–92.
Stockwell, R. – J. Bowen
1965 *Sounds of English and Spanish*. Chicago: University of Chicago Press.
Strange, W.
1989 "Dynamic specification of coarticulated vowels spoken in sentence context", *Journal of the Acoustical Society of America* 85:2135–2153.
Tahta, S. – M. Wood – K. Loewenthal
1981 "Foreign accents: Factors relating to transfer of accent from the first language to a second language", *Language and Speech* 24:265–272.
Trubetzkoy, N.
1939/1969 Grundzüge der Phonologie. Travaux du Cercle Linguistique de Prague 7. (English translation by C. Baltaxe) Berkeley California: University of California Press.
van Heuven, V.
1986 Personal communication.
Weinreich, U.
1953 *Languages in Contact: Findings and Problems*. The Hague: Mouton.
Wells, J.
1982 *Accents of English*, Vol. 1, "An Introduction". Cambridge: Cambridge University Press.
Wode, H.
1977 "The L2 acquisition of /r/", *Phonetica* 34:200–217.
Wode, H.
1978 "The beginning of non-school room L2 phonological acquisition", *International Review of Applied Linguistics* 16:109–125.

Perpeption and production of a new vowel category by adult second language learners[1]

Ocke-Schwen Bohn – James Emil Flege

1 Introduction

This paper addresses two questions in L2 speech research: can adults learn to produce and perceive a second language vowel category for which no counterpart exists in their native language? And, secondly, if such learning of an L2 vowel by adults can be observed, what is the relation between their production and perception of the new category?

The first question is of interest not just in L2 speech research, but in L2 research in general. The still influential critical period hypothesis states that adults cannot overcome biologically conditioned limits on the ability to learn (an) additional language(s) successfully. With respect to L2 speech learning, this hypothesis would predict that adults will not establish phonetic categories for sounds which are not found in their L2. This view has recently been challenged by Flege (1987), who showed that adults can produce foreign vowels authentically if these vowels are unlike any native vowel, and if as learners they have had extensive exposure to the foreign language. Flege hypothesized that vowels which are sufficiently different from native vowels (like, in his study, French /y/ for native English speakers) will not be treated as equivalent to existing phonetic categories by L2 learners, and that sufficient input will enable adult learners to establish phonetic categories for new vowels.

The present study further tested the hypothesis that adult learners will eventually establish new phonetic categories for vowels that do not have an easily identifiable counterpart in the native language. It differs from previous studies of second language speech in that it considers both the perception and the production of a new vowel category by two groups of non-native speakers who differed in L2 experience. This made it possible to study the effect of L2 experience on category formation for new vowels, and to study the relation between perception and production in the acquisition of a new vowel category.

The results to be presented were obtained in a larger study which investigated the perception and the production of several English vowels, including

the vowel /æ/, by native speakers of German (see Bohn–Flege 1989; 1990; 1992). Perception of this new vowel for native Germans was tested in an identification experiment in which a synthetic *bet-bat* continuum was presented to two groups of native German listeners differing in experience with English. Production of English /æ/ by the same two German groups was assessed from recordings of bVt words, including *bat*. In both the production and the perception experiment, the same monolingual English control group participated.

An essential preliminary step in our study was to establish whether English /æ/ is indeed a new vowel for native Germans. Previous research comparing English and German vowels suggests that this is indeed the case if Standard German is the native dialect. There are, however, German dialects which have /æ/ as part of their phonemic and phonetic repertoire. Since the hypothesis to be tested in the present study depended crucially on whether English /æ/ *is* a new vowel for native Germans, the first production experiment reported here aimed to establish this relationship for the subjects in our study.

The two groups of native Germans who participated in the production and perception experiments that focused on English /æ/ also read lists of German sentences (*Ich sage* /bVt(ə)n/) containing the German vowels /ɛ/, /a/, and /ɛ:/ or /e/ in the words *betten, batten,* and *bäten,* respectively. (Orthographic <ä> in the last word is either /ɛ:/ or /e/, depending on the speaker's regional origin and speaking style.) The German vowels in these words were chosen because they are located in the same general area of the vowel space as English /æ/. The comparison of German /ɛ/, /ɛ:/~/e/, and /a/ to English /æ/ is based on acoustic analyses of five repetitions of each vowel by 10 monolingual English speakers for /æ/, and by 10 native German speakers (the group less experienced in English) for the three German vowels. Measurements using linear predictive coding (LPC) analysis included the fundamental and F_1, F_2, and F_3 at the acoustic midpoint of the vowel. The frequency values were converted into Barks to obtain Bark-difference scores (Syrdal 1985; Syrdal–Gopal 1986), which normalized the gender differences in our data.

Figure 1 compares the position of English /æ/ to the position of the three German vowels from the same area of the vowel space. The results are presented in the perceptually valid Bark-difference space. The axes in Figure 1 represent the front-back ($Bark_2$-$Bark_1$) and vowel height ($Bark_1$-$Bark_0$) dimensions of traditional vowel diagrams. The plot clearly shows that the German subjects in this study have no vowel in their native language that occupies the area taken up by English /æ/. This strongly suggests that for the native Germans in this study, English /æ/ is a new vowel.

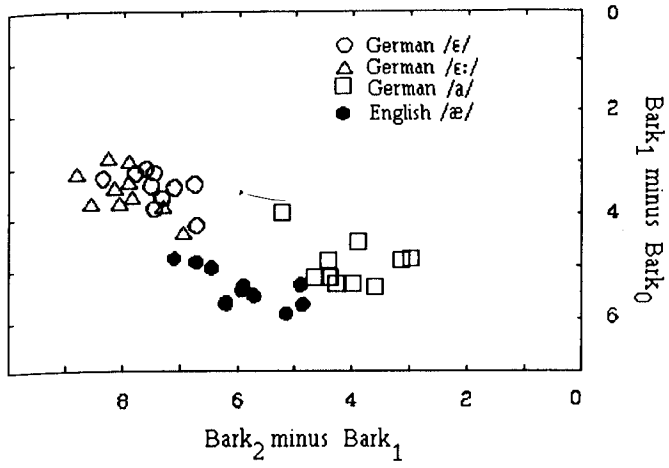

Figure 1. Distribution of the English vowel /æ/ (as produced by monolingual native English speakers) and the German vowels /ɛ/,/ɛ:/~/e/, and /a/ (as produced by native German speakers with relatively little English language experience) in the Bark-difference space.

Having established that the English /æ/ vowel has acoustic properties that are not shared by any German vowel, the specific hypothesis derived from Flege's speech learning model (Flege 1987; 1988; 1995; see also Flege–Hillenbrand 1984) is that the experienced, but not the inexperienced native German speakers of English will closely match the native English speakers (a) for their productions of the new English vowel /æ/, and (b) for perception of /æ/, which was tested here in an identification experiment.

2 Production of the new vowel /æ/

The acoustic data reflecting the production of the new vowel /æ/ were obtained from two groups of native Germans differing in English language experience, and from a monolingual native English group. The experienced Germans had lived in the United States for at least five years (mean: 7.5 years), whereas the inexperienced Germans had recently arrived (mean length of stay in the US: 0.6 years). The mean age of the experienced and inexperienced Germans was 28 and 33 years, respectively. They had studied English in school for about the same number of years (7.6 and 6.6 years, respectively). The native English control group consisted of monolingual

subjects (mean age 28 years). The subjects in the three groups participated as paid volunteers.

Each subject produced English consonant-vowel-consonant words in the carrier phrase *I will say___*. The subjects read counterbalanced blocks of randomized sentences in which the target vowel was held constant. The results reported here are based on 5 repetitions by each subject of English *bat*, which contains the new vowel /æ/, and of *bet*, which contains the neighboring vowel /ɛ/.

The duration and frequencies of the fundamental (F_0) and the first three formants (F_1, F_2, F_3) were analyzed for each vowel. As for the comparison of English /æ/ to neighboring German vowels, the frequency values were converted into Barks to obtain Bark-difference scores, which normalized the gender differences in our data. Following Syrdal and Gopal (Syrdal 1985; Syrdal–Gopal 1986), we will refer to the $Bark_1$-$Bark_0$ dimension as the vowel height dimension, and to $Bark_2$-$Bark_1$ as the front-back dimension.

Figure 2 compares the intended productions of English /æ/ and the neighboring English vowel /ɛ/ by the native English (top panel), the experienced (center panel) and the inexperienced German group (bottom panel) in the Bark-difference space.

Figure 2 shows a fairly clear separation of the /ɛ/ and /æ/ categories for the native English and the experienced German subjects, and an almost complete overlap for the inexperienced German subjects, indicating that the experienced, but not the inexperienced Germans produced an English-like /ɛ-æ/ contrast.

Statistical analyses supported this conclusion. Separate one-way ANOVAs testing the effect of Group (native English vs experienced vs inexperienced Germans) on the Bark-difference scores revealed a significant Group effect for the vowel height dimension (B_1-B_0: $F(2, 27) = 6.653$, $p < .01$), but not for the front-back dimension (B_2-B_1: $F(2, 27) = 1.17$, $p > .05$). The effect for B_1-B_0 was obtained because the mean score for the inexperienced Germans (4.50) was significantly smaller than the score for the native English group (5.35) ($p < .01$, Newman-Keuls) and the inexperienced German group (5.07) ($p < .05$, Newman-Keuls). The experienced Germans did not differ significantly from the native English speakers ($p > .05$, Newman-Keuls). This suggests that the intended /æ/s of the inexperienced Germans were higher in the acoustic vowel space than the /æ/s produced by the experienced Germans and the native English group.

The duration of /æ/ as produced by the three groups were not compared directly, but based on each speaker's duration ratio for English /æ/ to its neighboring vowel /ɛ/. Figure 3 presents individual subject duration ratios (/æ/-duration divided by /ɛ/-duration) for speakers in the three groups.

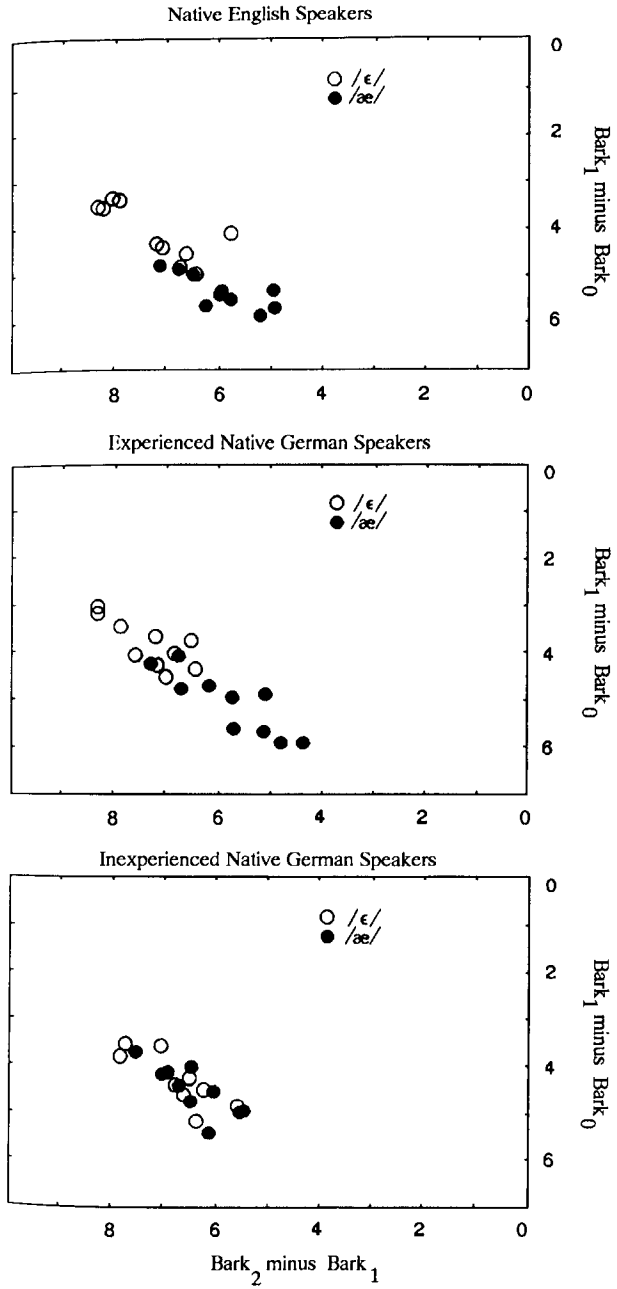

Figure 2. Distribution of the English vowels /ɛ/ and /æ/ as produced by the native English group (top panel), the experienced German group (center panel), and the inexperienced German group (bottom panel) in the Bark-difference space.

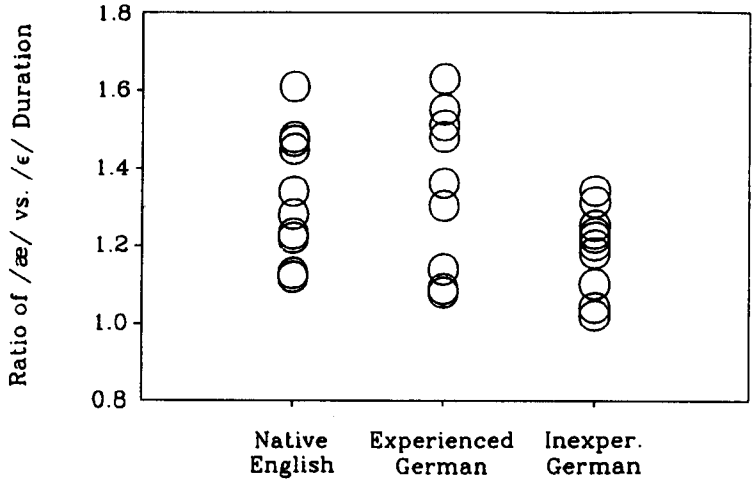

Figure 3. Duration ratios (/æ/-duration divided by /ɛ/-duration) for individual speakers in the three groups.

The range of ratios, and the distribution within the ranges, is fairly similar for the native English and the experienced German speakers. The ratios for the inexperienced Germans, however, cluster around smaller values, indicating that, on average, members of this group produced a smaller duration contrast between /æ/ and /ɛ/ than the experienced Germans and the native English speakers.

A one-way ANOVA testing the effect of Group on the /æ-ɛ/ duration ratios revealed a barely significant Group effect ($F(2,27) = 3.352$, $p = .0496$). Newman-Keuls post-hoc tests did not reach the 5% significance level for any between-group comparison. Pairwise comparisons for the three groups in one-way ANOVAs studying the effect of Group (2 levels) on /æ/-/ɛ/ duration ratios revealed that the ratio for the inexperienced Germans (1.2) was significantly smaller than for the native English (1.3) ($F(1,19) = 5.27$, $p < .05$) and the experienced German group (1.3) ($F(1,19) = 5.843$) $p < .05$). The native English and the experienced German group did not differ significantly from one another ($F(1,18) = .117$, $p > .05$).

These results support the hypothesis that experienced, but not inexperienced adult learners will produce a new vowel authentically. As far as the acoustic properties measured are concerned, the experienced Germans did not differ from the native English subjects in producing the new vowel /æ/. The intended /æ/ of the inexperienced Germans, however, differed from both the native English and the experienced German group in terms of spectral properties and in terms of the duration contrast between the two spectrally neighboring vowels /ɛ/ and /æ/: It was higher in the acoustic vowel space than the other groups' /æ/ and it was relatively short, which made the intended /æ/ productions of the inexperienced Germans very similar to their intended /ɛ/ productions.

3 Perception of the new vowel /æ/

A perception experiment was carried out to determine how labeling responses to a continuum involving the new vowel /æ/ would relate to the findings from acoustic measurements, which suggested that foreign language experience enables learners to establish a new phonetic category.

The same subjects as in the production experiment participated. A continuum ranging from *bet* to *bat* was created using the parallel mode of the Klatt (1980) software synthesizer. Two parameters were varied factorially – duration in three linearly equal steps and vowel spectrum in eleven linearly equal steps – generating 33 stimuli for the *bet-bat* continuum. The formant frequencies for F_1–F_3 varied from values appropriate for English /ɛ/ to those for English /æ/[2]. The vowel portion for each of the 11 spectrally different stimuli had nominal durations of 150, 200, and 250 ms[3].

In a self-paced experiment the native English and the experienced and inexperienced native German subjects listened individually to eleven randomizations of the stimuli and identified them as *bet* or *bat* by pressing one of two buttons of a response box. The results, which are based on responses to all but the first of the eleven randomizations, were tabulated in terms of per cent *bet* responses.

Figure 4 presents the group identification functions for the *bet-bat* continuum. The panels show per cent *bet* responses to each of the 11 spectrally differing stimuli. Stimulus number 1 refers to the /ɛ/ endpoint, and stimulus number 11 to the /æ/ endpoint of the continuum. As expected, the native English subjects (top panel) showed a clear crossover from *bet* to *bat* judgments as vowel spectrum varied between the endpoint stimuli, which were unambiguously identified. The identification functions for stimuli of short, medium, and long duration are on top of each other except for those stimuli that were spectrally ambiguous between *bet* and *bat*. This indicates that differences in vowel duration had little influence on how the native English listeners distinguished between /ɛ/ and /æ/.

The center panel shows how the inexperienced German listeners labeled stimuli from the *bet* to *bat* continuum. Unlike the native English listeners, the

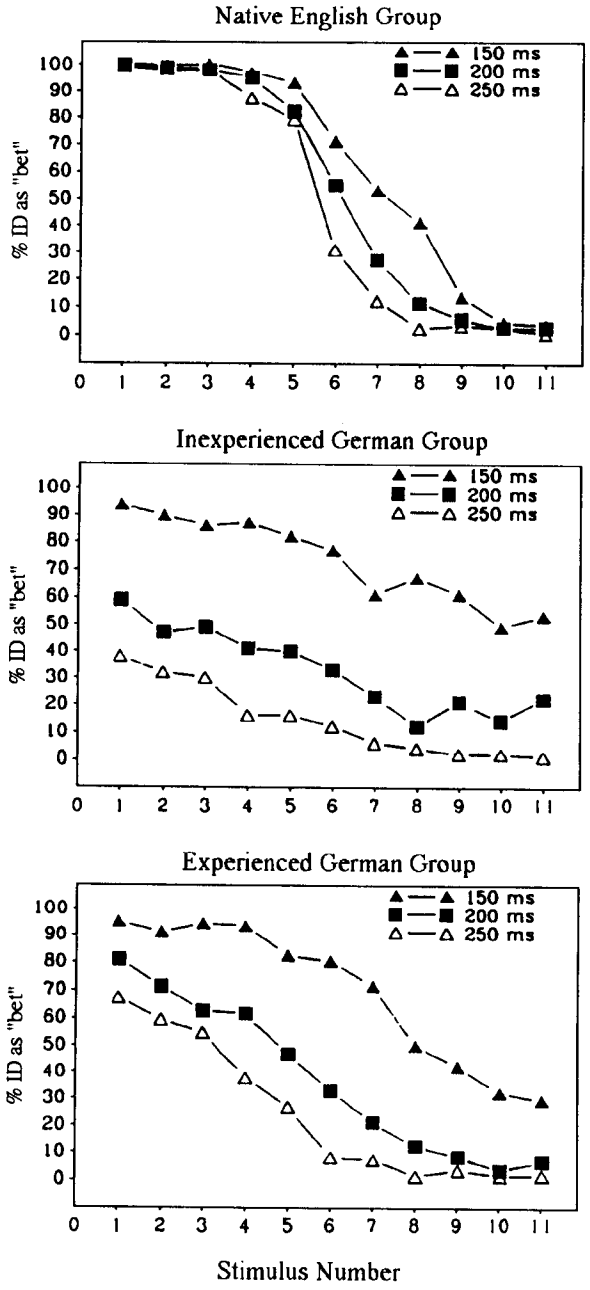

Figure 4. Group identification functions of the *bet-bat* continuum. (Stimulus number 1 refers to the spectral /ɛ/ endpoint and stimulus number 11 to the spectral /æ/ endpoint of the continuum.)

inexperienced Germans were strongly influenced by duration in labeling stimuli as *bet* or *bat*. They identified short stimuli mostly as *bet* and long stimuli mostly as *bat*. Stimuli of medium duration were judged to be ambiguous. The inexperienced Germans did not show clear crossovers for any of the three vowel durations, and they identified unambiguously only two of the six spectral endpoint stimuli, namely the short /ɛ/ endpoint and the long /æ/ endpoint . However, vowel spectrum did influence the labeling of stimuli as *bet* or *bat* somewhat. This can be seen from the shallow and fairly monotonic slopes of the identification functions, which show consistently higher percentages of *bet* judgments for stimuli near the /ɛ/ than for those near the /æ/ endpoint.

The bottom panel presents the identification functions for the experienced German listeners. These functions show that the responses of the experienced Germans were more influenced by spectral differences than those of the inexperienced Germans. The experienced Germans showed crossovers from predominantly *bet* to predominantly *bat* judgments for all three durations. However, the experienced Germans also differed from the native English listeners. They were more influenced by duration differences and less by spectral influences than the native English listeners. The identification functions of the experienced Germans for stimuli of short, medium, and long duration are not on top of one another. Rather, there was a clear tendency for shorter stimuli to be labeled as *bet* and longer stimuli to be labeled as *bat*. The overall impression from Figure 4 is that, in their labeling of stimuli from the *bet-bat* continuum, the experienced Germans listeners were more similar to the native English than the inexperienced Germans. This conclusion is supported by various analyses of the responses.

One simple way to compare the group responses is based on a count of spectrally-based crossovers in *bet* vs *bat* judgments for stimuli of short, medium, or long vowel duration. A crossover was said to occur if the difference between the responses to the endpoint stimuli (1 and 2 vs 10 and 11) was 70 % or greater. The application of this criterion is illustrated by Figure 5, which presents the identification functions for two German subjects from the experienced group.

One (top panel) showed a sharp crossover from *bet* to *bat* responses as the spectrum changed. This subject met the crossover criterion for all three vowel durations. The other subject (bottom panel) based her responses primarily on vowel duration and did not meet the criterion for spectrally-based crossovers for any of the three durations.

The total number of responses to stimuli of short, medium, or long vowel duration that met the crossover criterion was 18 for the experienced and 6

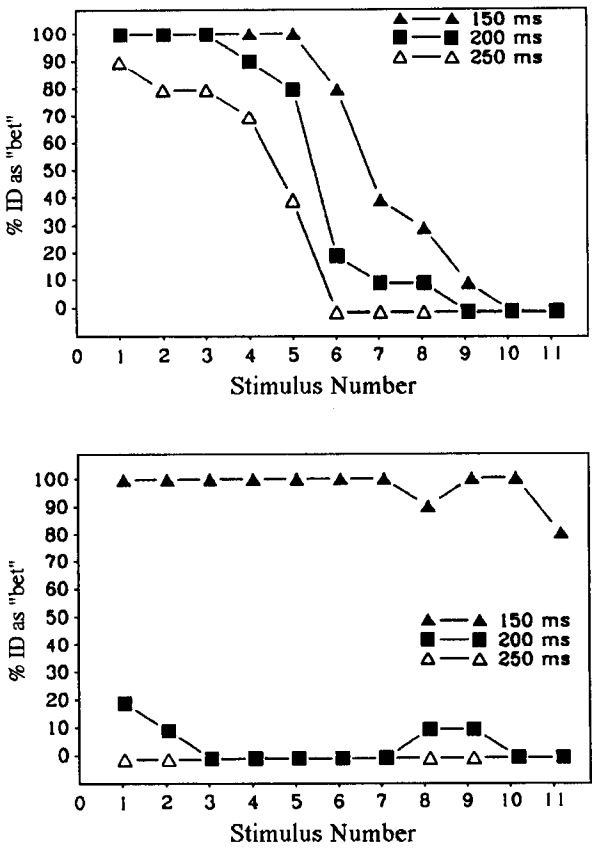

Figure 5. Identification functions for the *bet-bat* continuum for two experienced native German speakers of English (see text).

for the inexperienced Germans, as compared to the maximum possible number of 30 for the native English listeners. The fact that more crossovers were observed for the experienced than the inexperienced Germans suggests again that English language experience may cause native German listeners' perception of the /ɛ/-/æ/ contrast to become more English-like.

Responses from subjects in the three groups were further analyzed by submitting the per cent *bet* responses to the 33 stimuli to a three-way ANOVA with Group (3 levels) as between-subjects factor and Duration (3 levels) and Spectral Step (11 levels) as within-subjects factors. A significant Group × Duration × Spectrum interaction was obtained ($F(40,540) = 2.223$, $p < .001$) in addition to significant interactions for Group x Duration ($F(4,54) = 7.300$, $p < .001$) and Group × Spectrum ($F(20,270) = 11.487$, $p < .001$). The three-

way interaction was explored through separate ANOVAs for each duration (short, medium, long) testing the effect of Group on the change in *bet* responses between the spectral endpoint stimuli. The response changes due to the spectral manipulation were significant for each group in all three durations ($p < .01$). For stimuli with short vowel duration the change in labeling between the spectral endpoint stimuli was smaller for the inexperienced Germans (41%) than the native English listeners (97%), whereas the change for the experienced Germans (67%) did not differ significantly from the other two groups (Newman-Keuls, $p < .05$). For stimuli with medium vowel duration the effect of the spectral manipulation was smaller for inexperienced Germans (37%) than for the native English (98%) and the experienced German (76%) groups, who did not differ significantly from each other (Newman-Keuls, $p < .05$). For stimuli with long vowel duration the change in *bet* responses as a result of the spectral manipulation was larger for the native English listeners (99%) than for both German groups (GB: 67%, GA: 31%), who did not differ significantly from each other (Newman-Keuls, $p < .05$).

These results show that the effect of spectral manipulation was significantly smaller for the inexperienced Germans than for the native English group for all three durations, whereas for two durations (short and medium) this effect did not differ significantly between the native English and the experienced German listeners. The conclusion that prolonged exposure to English leads native Germans to perceive the English /ɛ/-/æ/ contrast in a more native-like manner receives further support from the exploration of the Group × Duration interaction. The response changes resulting from the duration manipulation, which were significant for all three groups, ($p < .001$), were smaller for the native English (15%) than for the experienced German group (45%), who in turn showed a smaller duration effect than the inexperienced Germans (59%).

These findings indicate that the three groups differed with respect to the relative effect of spectral and duration manipulations in judging vowels as /ɛ/ or /æ/. The spectral cues were more important for the native English than the experienced Germans than the inexperienced Germans, and conversely, the duration cue was more important for the inexperienced than the experienced Germans – for whom, however, it was more important than for the native English listeners. This suggests that English language experience affects how native German listeners perceive English /ɛ/ and /æ/.

To conclude, the results of the identification experiment showed that a larger number of experienced than inexperienced German subjects identified stimuli from a continuum that involved the new English vowel /æ/ in a way

that was similar to the native English listeners. That is, more Germans from the experienced than from the inexperienced group relied on spectral rather than temporal cues – suggesting that extended contact with English may precipitate an English-like perception of the /ɛ/ vs /æ/ contrast.

4 Relation between perception and production of a new vowel category

Both the production and the perception experiment provided support for the hypothesis that adults can establish a new vowel category, given sufficiently extensive second language input. The two experiments differed, however, in the strength of support they provide for the hypothesis. In the production experiment, the experienced German speakers of English did not differ significantly from the native English speakers in terms of the spectral and duration properties of the new vowel /æ/. In the perception experiment, however, the experienced Germans did not fully match the native English speakers. While the experienced Germans were clearly more English-like in their labeling responses than the inexperienced Germans, they did not make as much use of vowel spectrum, and as little use of duration, as the native English listeners in distinguishing between /ɛ/ and /æ/. Comparison of the perception with the production results for the experienced German group therefore suggests that second language experience has a more profound impact on the production than on the perception of the new vowel category.

To examine the relationship between production and perception of the new category in more detail, we compared (a) the use of the duration cue in the identification experiment with the duration ratio of /æ/ to /ɛ/ in production, and (b) the use of the spectral cues in the identification experiment with the difference in vowel height ($Bark_1$-$Bark_0$) between /ɛ/ and /æ/ productions. (The front-back dimension ($Bark_2$-$Bark_1$) was not included in the comparisons because the three groups did not differ significantly along this dimension; see above).

4.1 The relation between perception and production of spectral differences

Figure 6 presents individual subject data comparing the use of spectral cues in the identification experiment with the difference in vowel height ($Bark_1$-$Bark_0$) between /ɛ/ and /æ/ productions.

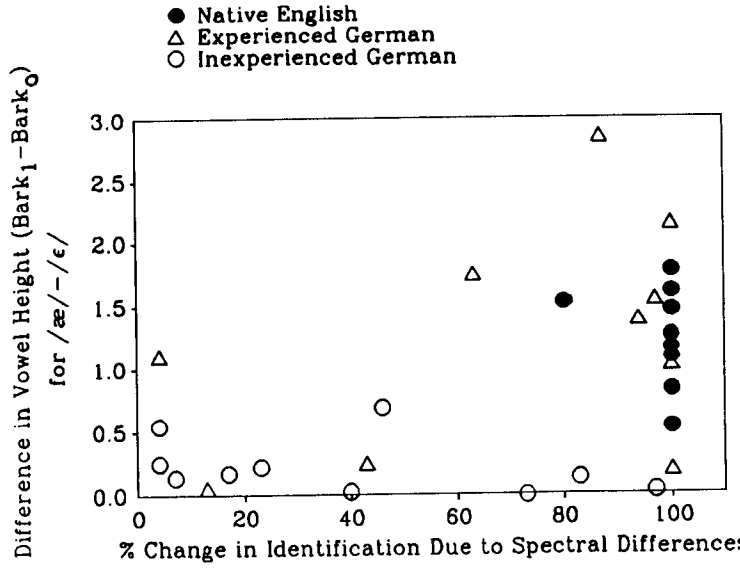

Figure 6. Plot of individual subjects' use of spectral cues in the identification experiment vs difference in vowel height ($Bark_1$-$Bark_0$) between /ɛ/ and /æ/ in production.

The plot shows that a strong influence of spectral differences on labeling responses did not correspond to a large spectral difference between /ɛ/ and /æ/ in production, and that those subjects who produced a relatively small spectral contrast differed greatly in their use of spectral cues in the identification experiment. The absence of data points in the upper left portion of the plot indicates that no subject in any of the three groups showed a small influence of spectral differences on perception and produced a large spectral difference between /ɛ/ and /æ/.

While the native English listeners showed a uniformly high effect of spectral differences in their labeling of endpoint stimuli as *bet* or *bat*, they were quite variable with respect to the magnitude of the spectral difference between /ɛ/ and /æ/ in production. Conversely, the inexperienced German uniformly produced no spectral difference or only a small one between /ɛ/ and /æ/, but were quite variable in the use of spectral cues in the identification experiment. The results for the experienced Germans were intermediate between the results for the other two groups: those subjects who were strongly influenced by spectral differences in the identification experiment differed greatly in the magnitude of spectral differences between /ɛ/ and /æ/ in production, and those subjects who produced no spectral difference,

or only a small one, differed greatly in their use of spectral cues in perception.

We may conclude that, for all three groups, production and perception of spectral differences between /ɛ/ and /æ/ were quite independent of each other. The important reservation is that relatively large spectral differences in production imply a strong influence of spectral cues in perception as illustrated by the empty upper left portion of Figure 6.

4.2 The relation between perception and production of duration differences

Figure 7 presents individual subject data comparing the use of the duration cue in the identification experiment with the use of duration in differentiating /ɛ/ from /æ/ in production.

The plot shows that a strong reliance on duration in differentiating /ɛ/ from /æ/ in identification does not correspond to a large duration contrast between these vowels in production, and that those subjects who produced a large duration contrast were not particularly susceptible to duration changes in their labeling responses. The absence of data points from the upper right

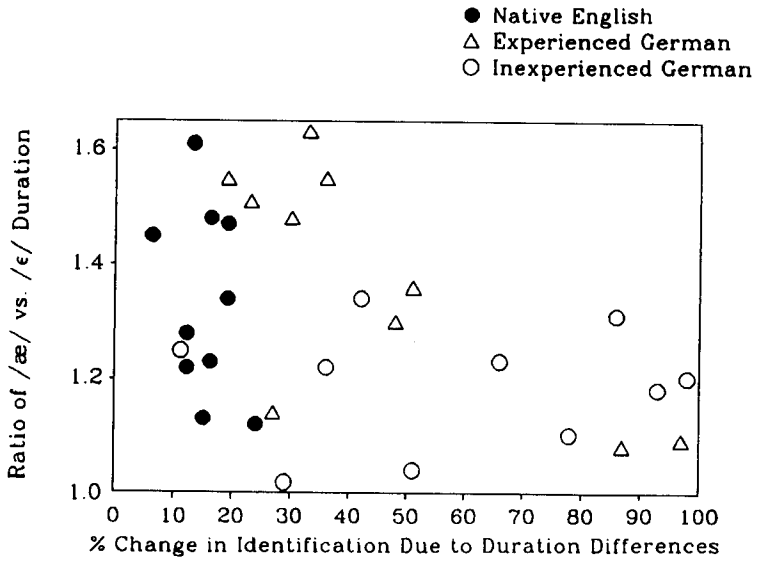

Figure 7. Plot of individual subjects' use of the duration cue in the identification experiment vs duration ratio of /æ/ to /ɛ/ in production.

portion of the plot is noteworthy because it indicates that no subject in any of the three groups produced a large duration difference between /ɛ/ and /æ/ and was strongly influenced by duration differences in perception.

The native English listeners did not make much use of the duration cue in identification, and members of this group varied little in their use of this cue. In production, however, some native English speakers produced fairly large and others quite small duration contrasts between /ɛ/ and /æ/. The inexperienced Germans present a picture opposite to that obtained for the native English group: they varied little in producing relatively small duration differences, but differed greatly among themselves in their use of the duration cue in perception. Most of the experienced Germans were not strongly influenced by duration differences in the identification experiment, and those that made least use of this cue tended to produce large duration differences between /ɛ/ and /æ/.

The conclusion from this comparison is similar to the one drawn for spectral differences: for all three listener groups production and perception of duration differences between /ɛ/ and /æ/ are quite independent of each other. Again, there is one major reservation, namely that a strong influence of duration differences in perception implies a small duration contrast in production.

5 Conclusion

Three major findings have resulted from this study of the perception and production of a new vowel category by adult L2 learners. First, given extensive foreign language experience, adults can learn to produce and perceive a new vowel category in a way similar to native speakers of the L2. Second, L2 experience seems to have a more profound impact on the production than on the perception of a new vowel category. Third, for both native speakers and L2 learners, perception and production of spectral and duration differences for the English /ɛ/-/æ/ pair were quite independent of each other, except that different implicational relations seemed to exist between production and perception for spectral and for duration differences.

The first finding is consistent with Flege's (1987; 1988; 1995) speech learning model, which states that L2 sounds for which no obvious counterparts exist in the L1 will evade equivalence classification, so that the L2 learners will eventually establish phonetic categories for new L2 sounds. The results of this study point to the importance of cross-language sound correspondences, or lack thereof, in L2 speech learning, and they strongly suggest that a

critical period for learning new sounds does not exist, at least not before the age of ca. 30 years (the mean age of our subjects).

The second finding adds to the growing evidence from L2 speech research which indicates that production of L2 sounds is easier and/or faster to learn than perception of the same sounds (e.g., Brière 1966; Goto 1971; Caramazza et al. 1973; Williams 1977; Garnes 1978; Mochizuki 1981; Sheldon–Strange 1982; Tees–Werker 1984; Gass 1984; Sheldon 1985; Flege–Eefting 1987; Hammarberg 1988). According to other studies, however, production of L2 sounds may lag behind their perception (e.g., Oakshott–Taylor 1976; Barry 1977; Neufeld 1988; Flege–Hillenbrand 1984).

The complex issue of perceptuo-productive heteromorphism in L2 speech learning cannot be discussed exhaustively in this chapter, one reason being that it is not at all clear how perception and production are linked in mature native speakers[4]. However, it appears likely that the following approaches might help clarify the relation between production and perception in L2 speech learning. One way would be to consider in detail each relevant study's criteria for "easier to learn", "accuracy", etc. in production and in perception. The criteria used to evaluate learners' performance and the reliability of accuracy judgments may vary considerably from study to study (see Strange–Broen 1980 and Leather 1983 for reviews). Second, the possibility should be considered that the relation between production and perception may not be the same for different classes of sounds, such as consonants and vowels. Third, the relation may differ according to whether the sounds of the L2 are similar to L1 sounds, or do not have a counterpart in the L1 (as in the present study).

While the results of the present study do not afford an opportunity to test these approaches, Sheldon's (1985) attempt to clarify the relation between production and perception in L2 speech learning can be fruitfully applied to the present study. Sheldon's reanalysis of the data in the Borden et al. (1983) study on the perception and production of L2 English /r/ and /l/ by L1 Korean speakers led her to conclude that "perception and production abilities interrelate in different ways during the acquisition process" (Sheldon 1985: 111) and that the relationship appeared to depend on the amount of time the learner had spent in the United States.

The same conclusion can be drawn for the present study. The inexperienced Germans failed to produce an /ɛ/-/æ/ contrast, but differentiated the two vowels systematically in the labeling experiment. However, they relied primarily on a cue (duration) that does not strongly influence native speakers of English. The experienced Germans, on the other hand, produced a native-like /ɛ/-/æ/ contrast, but still did not use spectral information as much as the native English listeners in differentiating /ɛ/ and /æ/. This suggests that inexperienced

L2 learners may differentiate a new vowel contrast perceptually (albeit on the basis of a non-native-like criterion) without differentiating this contrast in production. In the early stages of L2 speech learning, perception may therefore lead production, although the perceptual criteria may be very different from those used by native speakers. The results for the experienced Germans suggest that continued L2 contact enables L2 speakers to produce a new vowel contrast like native speakers of the L2, and that perception abilities for a new vowel contrast may lag behind even after several years of L2 experience.

One can only speculate why experience may have a more profound impact on the production than on the perception of L2 sounds. Perhaps perception of a new vowel contrast is more resistant to L2 experience than production because speech production is more subject to social control than speech perception. Non-native listeners can function adequately if they rely on vowel spectrum and duration to differentiate the English /ɛ/-/æ/ contrast because native English speakers produce both spectral and duration differences between /ɛ/ and /æ/. On the other hand, non-native talkers may feel greater pressure to conform to the production norms of the L2 in order to avoid being stigmatized for misidentifiable or foreign-accented speech.

The finding that L2 perception and production abilities do not progress in parallel has led some authors to conclude that inferences about L2 production cannot be drawn from L2 perception, and vice versa (e.g. Barry 1974; Weiher 1975; Jansma 1987). The third major finding of the present study provides support for this view. Perception and production of spectral and duration differences for the English /ɛ/-/æ/ pair were quite independent of each other, no matter whether the subjects were native speakers or L2 learners differing in L2 experience. However, an interesting finding from the separate perception/production comparisons for spectral and duration differences was that different relations were revealed between perception and production. A positive link seemed to exist between the production and the perception of spectral differences, since all subjects who produced large spectral differences were also strongly influenced by spectral cues in perception. For duration differences, however, a negative link was observed: all subjects who produced a large duration contrast between /ɛ/ and /æ/ were relatively insensitive to duration differences in perception. These findings suggest that different relations between production and perception may exist for different acoustic correlates (such as vowel spectrum and duration) of phonetic categories.

Notes

1. Research supported by grant NS20963-04 from the U.S. National Institutes of Health and conducted at the Department of Biocommunication at the University of Alabama in Birmingham, USA. Part of the results reported here were presented at the Winter Meeting of the Linguistic Society of America, New Orleans, Louisiana, USA, 27.–30. December 1988, and at the 117th Meeting of the Acoustical Society of America, Syracuse, New York, USA, 22–26 May 1989.
2. Peterson–Barney (1952) male values were used for stimuli 2 and 10. The values for the endpoint stimuli 1 and 11 were extrapolated. The nominal values for F_1–F_3 in the /ɛ/ endpoint stimulus were: 515, 1830 and 2490 Hz. The /æ/ endpoint had nominal F_1–F_3 values of 675, 1730 and 2400 Hz.
3. The actual durations as measured with a waveform editor were 138, 190, and 233 ms. These durations included the formant transitions used to cue the word-initial and -final stops.
4. For instance, Fox (1978; 1982) and Bell-Berti et al. (1979) reported a clear link between the perception and production of vowels for their adult monolingual subjects, whereas Paliwal et al. (1983) and Gottfried and Beddor (1988) found that their listeners were not referring to their own vowel productions for vowel perception. The results of the present study for the monolingual English listeners are very similar to those reported by Gottfried and Beddor for native French speakers. The subjects in both studies were quite insensitive to duration differences in perceptually distinguishing native vowel pairs (/ɔ/-/o/ in the Gottfried and Beddor study, /ɛ/-/æ/ in the present study), but they did produce a temporal contrast between /ɔ/-/o/ and /ɛ/-/æ/, respectively.

References

Barry, William J.
1974 *Perzeption und Produktion im subphonemischen Bereich. Eine kontrastive Untersuchung an intersprachlichen Minimalpaaren des Deutschen und Englischen.* Tübingen: Niemeyer.

Barry, William J.
1977 "Ausspracheabweichungen: Eine experimentalphonetische Untersuchung an deutschen Englischlernenden", Institut für Phonetik der Universität Kiel, *Arbeitsberichte* 7.

Bell-Berti, Fredericka – Raphael, Lawrence J. – Pisoni, David B. – Sawusch, James R.
1979 "Some relationships between speech production and perception", *Phonetica* 36:373–383.

Bohn, Ocke-Schwen – Flege, James Emil
1989 "The establishment of a new phonetic category by German learners of English", *Journal of the Acoustical Society of America* 85:S84–S85.

Bohn, Ocke-Schwen – Flege, James Emil
1990 "Interlingual identification and the role of foreign language experience in L2 vowel perception", *Applied Psycholinguistics* 11:303–328.

Bohn, Ocke-Schwen – Flege, James Emil
1992 The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14:131–158.

Borden, Gloria – Gerber, Adele – Milsark, Gary
1983 "Production and perception of the /r/-/l/ contrast in Korean adults learning English", *Language Learning* 33:499–526.

Brière, Eugene J.
1966 "An investigation of phonological interference", *Language* 42: 768–796.

Caramazza, A. – Yeni-Komshian, G.H. – Zurif, E.B. – Carbone, E.
1973 "The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals", *Journal of the Acoustical Society of America* 54:421–428.

Flege, James Emil
1987 "The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification", *Journal of Phonetics* 15:47–65.

Flege, James Emil
1988 "The production and perception of foreign languages", in: Harris Winitz (ed.), *Human Communication and its Disorders.* Norwood NJ: Ablex, 224–401.

Flege, James Emil
1995 Second language speech learning: Theory, findings, and problems. In: Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research.* Baltimore: York Press, 233–277.

Flege, James Emil – Eefting, Wieke
1987 "Cross-language switching in stop consonant perception and production by Dutch speakers of English", *Speech Communication* 6:185–202.

Flege, James Emil–Hillenbrand, James
1984 "Limits on phonetic accuracy in foreign language speech production", *Journal of the Acoustical Society of America* 76:708–721.

Fox, Robert A.
1978 "Individual perceptual variation and a perception/production link in vowels", in: Donka Farkas – Wesley M. Jacobsen – Karol Todrys (eds.), *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society.* Chicago IL: Chicago Linguistic Society, 98–107.

Fox, Robert A.
1981 "Individual variation in the perception of vowels: implications for a perception-production link", *Phonetica* 39:1–2.

Garnes, Sara
1978 "Some effects of bilingualism on perception", in: J. Weinstock (ed.), *The Nordic languages and linguistics* 3. Proceedings of the Third Interntional Conference of Nordic and General Linguistics. Austin TX: The University of Texas at Austin, 352–359.

Gass, Susan
1984 "Development of speech perception and speech production in adult second language learners", *Applied Psycholinguistics* 5: 51–74.
Goto, Hiromu
1971 "Auditory perception by normal Japanese adults of the sounds 'l' and 'r'", *Neuropsychologia* 9: 317–323.
Gottfried, Terry L. – Beddor, Patricia S.
1988 "Perception of temporal and spectral information in French vowels", *Language and Speech* 31:57–75.
Hammarberg, Björn
1988 *Studien zur Phonologie des Zweitsprachenerwerbs*. Stockholm: Almqvist and Wiksell.
Jansma, Kimberly
1987 "English intonation: Do learners hear the pattern?" Paper presented at the Annual Meeting of the Linguistic Society of America, San Francisco CA, December 1987.
Klatt, Dennis H.
1980 "Software for a cascade/parallel formant synthesizer", *Journal of the Acoustical Society of America* 67:971–995.
Leather, Jonathan
1983 "Second-language pronunciation learning and teaching", *Language Teaching* 16:198–219.
Mochizuki, Michiko
1981 "The identification of /r/ and /l/ in natural and synthesized speech", *Journal of Phonetics* 9:283–303.
Neufeld, Gerald G.
1988 "Phonological asymmetry in second-language learning and performance", Language Learning 38:531–559.
Oakeshott-Taylor, John
1976 "Perception and production of English sounds by German speaking subjects", *Hamburger Phonetische Beiträge* 17:163–199.
Paliwal, K.K. – Lindsay, D. – Ainsworth, W.A.
1983 "Correlation between production and perception of English vowels", *Journal of Phonetics* 11:77–83.
Peterson, Gordon E. – Barney, Harold L.
1952 "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America* 24:175–184.
Sheldon, Amy
1985 "The relationship between production and perception of the /r/-/l/ contrast in Korean adults learning English: a reply to Borden, Gerber, and Milsark", *Language Learning* 35:107–113.
Sheldon, Amy – Strange, Winifred
1982 "The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception", *Applied Psycholinguistics* 3:243–261.

Strange, Winifred – Broen, Patricia
1980 "Perception and production of approximant consonants by 3-year-olds: a first study", in: Grace H. Yeni-Komshian–James F. Kavanagh – Charles A. Ferguson (eds.), *Child Phonology Vol. 2: Perception*. New York: Academic Press, 117–154.
Syrdal, Ann K.
1985 "Aspects of a model of the auditory representation of American English vowels", *Speech Communication* 4:121–135.
Syrdal, Ann K. – Gopal, H.S.
1986 "A perceptual model of vowel recognition based on the auditory representation of American English vowels", *Journal of the Acoustical Society of America* 79:1086–1100.
Tees, Richard C. – Werker, Janet F.
1984 "Perceptual flexibility: maintenance or recovery of the ability to discriminate non-native speech sounds", *Canadian Journal of Psychology* 38: 57–590.
Weiher, Eckart
1975 *Lautwahrnehmung und Lautproduktion im Englischunterricht für Deutsche*. Doctoral dissertation, Institute for Phonetics, Kiel University.
Williams, Lee
1977 "The perception of stop consonant voicing by Spanish-English bilinguals", *Perception and Psychophysics* 21:289–297.