SSLA, 14, 131–158. Printed in the United States of America.

THE PRODUCTION OF NEW AND SIMILAR VOWELS BY ADULT GERMAN LEARNERS OF ENGLISH

Ocke-Schwen Bohn
Kiel University, Germany

James Emil Flege
University of Alabama at Birmingham

The study reported in this paper examined the effect of second language (L2) experience on the production of L2 vowels for which acoustic counterparts are either present or absent in the first language (L1). The hypothesis being tested was that amount of L2 experience would not affect L1 German speakers' production of the "similar" English vowels /i, I, ɛ/, whereas English language experience would enable L1 Germans to produce an English-like /æ/, which has no counterpart in German. The predictions were tested in two experiments that compared the production of English /i, I, ɛ, æ/ by two groups of L1 German speakers differing in English language experience and an L1 English control group. An acoustic experiment compared the three groups for spectral and temporal characteristics of the English vowels produced in /bVt/ words. The same tokens were assessed for intelligibility in a labeling experiment. The results of both experiments were largely consistent with the hypothesis. The experienced L2 speakers did not produce the similar English vowels /i, I, ɛ/ more intelligibly than the inexperienced L2 speakers, nor did experience have a positive effect on approximating the English acoustic norms for these similar vowels. The intelligibility results for the new vowel /æ/ did

This research was supported by a grant from the U.S. National Institutes of Health (NS20963-04) and conducted at the Department of Biocommunication, University of Alabama at Birmingham, USA. The authors thank Laurie Skelton for her help in conducting the listening experiment and two anonymous reviewers for their comments on a previous draft of this article.

Part of the results reported here were presented at the 117th Meeting of the Acoustical Society of America, Syracuse, NY, USA, May 22–26, 1989, and at the New Sounds 90 Conference, Amsterdam, The Netherlands, April 9–12, 1990.

Address reprint requests to Ocke-Schwen Bohn, Englisches Seminar, Christian-Albrechts-Universität, Olshausenstr. 40, 2300 Kiel 1, Germany.

not clearly support the model. However, the acoustic comparisons showed that the experienced but not the inexperienced L2 speakers produced the new vowel /æ/ in much the same way as the native English speakers.

This study addresses two basic questions in foreign language speech research: Can adults learn foreign language (L2) sounds? And, is their success or failure to do so explicable in terms of sound correspondences between the native language (L1) and the L2?

A fairly common belief is that adults cannot produce L2 sounds authentically (i.e., like native speakers of the L2). However, a recent study by Flege (1987) showed that adults *can* produce an L2 vowel authentically if it is sufficiently dissimilar from any L1 vowel and if the learners have had sufficient L2 exposure. Flege found that highly experienced native English speakers of L2 French produced French /y/ authentically, whereas their French /u/ remained English-accented. He attributed the difference between the two L2 vowels to their differing categorical status vis-à-vis English vowels. French /u/ has a close counterpart in English /u/, which leads L1 English speakers to treat French /u/ as being equivalent to English /u/. French /y/, however, has no easily identifiable counterpart in English. Flege concluded that new vowels (like /y/ for L1 English speakers) evade equivalence classification, and that sufficient input will enable adult learners to establish phonetic categories for new vowels.

Support for the hypothesis that extended L2 experience enables adults to learn new but not similar L2 vowels comes from a recent perceptual study by Bohn and Flege (1990). In that study, two groups of L1 Germans differing in L2 English experience and an L1 English group identified members from synthetic *beat–bit* and *bet–bat* continua in which vowel duration and spectrum varied factorily. For the *beat–bit* continuum, whose endpoints were vowels that are similar to German vowels (i.e., /i, I/), both German groups differed from the English group, but no difference existed between experienced and inexperienced L1 German speakers of English. However, for the *bet–bat* continuum, which involved one similar (/ɛ/) and one new (/æ/) vowel, the experienced Germans more closely resembled the L1 English group than did the inexperienced Germans. The perceptual results suggested that extended contact with English may precipitate the English-like perception of the /ɛ/–/æ/ contrast.

This study further tested the hypothesis that adult learners are able to establish phonetic categories for new L2 sounds, and thus eventually produce them authentically, whereas similar sounds will remain foreign-accented even after lengthy exposure to the L2 because category formation is blocked by equivalence classification. The experiments to be reported examined the production of English[1] /i, I, ɛ, æ/ by L1 German speakers. To test the hypothesis, it is first necessary to establish which sounds in a given L1/L2 combination are new, similar, or identical. The relationship of English /i, I, ɛ, æ/ to German vowels was explored through a literature review

and through an experiment examining acoustic properties of vowels produced by the L1 English and L1 German subjects of the present study (Experiment 1). Two experiments tested the predictions generated by the hypothesis through acoustic and perceptual measures. Experiment 2 compared spectral and duration properties of English /i, I, ɛ, æ/ as produced by experienced and inexperienced L1 German speakers and a monolingual native English group. Experiment 3 examined how native English listeners categorized the vowels examined acoustically in Experiment 2.

LITERATURE REVIEW OF THE RELATIONSHIP BETWEEN ENGLISH /i, I, ɛ, æ/ AND GERMAN VOWELS

For both theoretical and empirical reasons, the relationship among sounds of different languages cannot simply be stated by comparing phoneme inventories (Anderson, 1974; Disner, 1983; Kohler, 1971, 1981; Ladefoged, 1978). However, we may start by noting that both English and German have vowel phonemes that are usually transcribed as /i/, /I/, and /ɛ/. English also has an /æ/ phoneme, which is not found in most German dialects, including Standard German. Some descriptions of German include /ɛ:/ (e.g., Wurzel, 1984), which however is due to spelling pronunciation (Kohler, 1977; Moulton, 1987) and does not occur in colloquial Standard German, which has /e:/ instead.

Although the sounds of German and English have been examined thoroughly, acoustic cross-language comparisons have been rare. The high front vowels /i, I/ of English differ in terms of spectral quality and duration (House, 1961; Peterson & Lehiste, 1960). The primary difference for this pair is spectral, whereas the duration difference may be quite small (Barry, 1974a, 1974b; Delattre, 1964; House, 1961). The role of spectral versus duration differences in distinguishing German /i/–/I/ is less clear. Some acoustic studies support the belief (Jones, 1960; Moulton, 1962; Scherer & Wollmann, 1977) that German /i/ and /I/ are more peripheral than English /i/ and /I/ (Delattre, 1964; Jongman, Fourakis, & Sereno, 1989; but cf. Disner, 1983). The traditional view on the relative importance of duration versus spectral differences in German vowel pairs such as /i/–/I/ is that the duration contrast is relatively small, and the spectral contrast relatively large, in northern varieties of German, whereas the reverse is said to be true of southern varieties (Leskien & Brugman, 1882, p. 11). Although there is some limited support for this view (e.g., Iivonen, 1987), a straightforward relation does not seem to exist between the regional origin of German speakers and their primary use of either spectral or duration differences to contrast vowel pairs such as /i/–/I/ (Barry, 1974a, 1974b; Bethge, 1963; Iivonen, 1983, 1987; Weiss, 1972, 1976). The one certain finding seems to be that the /i/–/I/ duration contrast is larger in German than in English (Barry, 1974a, 1974b; Bethge, 1963; Delattre, 1964; House, 1961; Iivonen, 1987; Rietveld, 1975).

Not surprisingly, English /ɛ/ is not mentioned as a problem in textbooks for German learners of English. Previous acoustic research indicates that English /ɛ/ is either identical to German /ɛ/ or is somewhat lower in the acoustic-phonetic space and longer than German /ɛ/ (Delattre, 1964; House, 1961; Iivonen, 1987; Jørgensen,

1969; Narahara, Okamoto, & Shimoda, 1977; Peterson & Barney, 1952). For those German speakers who have /ɛ:/, this vowel may be an additional counterpart to English /ɛ/.[2]

Unlike /i, I, ɛ/, English /æ/ does not have a counterpart in most German dialects. It is worth noting, however, that /æ/ is part of the phonemic and phonetic repertoire of several German dialects, for instance, Swiss German (Boesch, 1957) and Bavarian (Traunmüller, 1982). An [æ] may also occur in North German dialects in words such as *Bert* [bæ(:)t], in which /r/ vocalization may result in a lowering and lengthening of the preceding vowel vis-à-vis its /r/-less counterpart (Bohn & Flege, 1990).

The literature review leads to the tentative conclusion that English /i, I, ɛ/ have close counterparts in German /i, I, ɛ, ɛ:/. Depending on the dialect background and/or idiosyncrasies of native German speakers, English /i, I, ɛ/ either are similar or may even be identical to German /i, I, ɛ, ɛ:/. English /æ/, however, has no readily identifiable counterpart in most German dialects and should therefore be considered a new vowel for native speakers of these dialects.

EXPERIMENT 1: ACOUSTIC COMPARISON OF ENGLISH /i, I, ɛ, æ/ AND GERMAN VOWELS

A speech production experiment was carried out to test the validity of these conclusions for the native German speakers whose English vowel productions were the focus of this study. Experiment 1 examined and compared acoustic properties of English /i, I, ɛ, æ/ as produced by native English speakers and of German /i, I, ɛ, ɛ:, a/ as produced by the native German speakers whose English productions were examined in Experiment 2 and Experiment 3.

Method

Subjects. Speech production data were collected from three groups of 10 adult subjects each (5 males, 5 females) who participated as paid volunteers. Two groups of L1 German speakers differed primarily in terms of English-language experience. Subjects in the "inexperienced" L1 German group (GA) had lived in an English-speaking environment for a mean of 0.6 years, whereas the mean for the "experienced" L1 German group (GB) was 7.5 years. The Germans had studied English in school for about the same number of years (GA: 6.6 years, GB: 7.6 years). The mean age was 28 years for GB and 33 years for GA.[3] The third group consisted of monolingual native English speakers (EN: mean age 28 years). Five of the 10 subjects in EN were natives of Alabama, and all EN subjects had spent several years of their adult lives in Birmingham, AL. Their productions probably provided reasonable representations of the native German speakers' L2 target. All of the subjects were affiliated with the University of Alabama at Birmingham at the time of the study.

Speech Materials. The German subjects produced five German words (*bieten, bitten, betten, batten, bäten*) in the carrier phrase *Ich sage* ____. The vowels in the

first four words are /i, I, ɛ, a/. The vowel in *bäten* may be /e:/ or /ɛ:/ (henceforth, /ɛ:/). The subjects in EN produced four English words (*beat, bit, bet, bat*), which contain the vowels /i, I, ɛ, æ/, in the carrier phrase *I will say* ____. The subjects read counterbalanced blocks of randomized sentences in which the target vowel was held constant. The four blocks for each English vowel consisted of 34 sentences, and the five blocks for each German vowel of 14 sentences, each. The English list contained five repetitions of *bVt* words and three repetitions of other consonant–vowel–consonant (CVC) words. The German list contained five repetitions of *bVt(ə)n* and three repetitions of other *CVC(ə)n* words.

The speech material was recorded in a single-walled sound booth using a high-quality cassette tape recorder (Marantz model PMD 420) with a stereo electret condenser microphone (Sony model ECM-939LT). The test words were low-pass filtered (4 kHz), digitized at 10 kHz with 12-bit amplitude resolution, and stored on disk for later analyses. The results to be reported are based on five repetitions by each German subject of German *bVt(ə)n* words and on five repetitions by each subject in EN of English *bVt* words.

Acoustic Measurements. The duration, fundamental frequency (F0), and frequency of the first three formants (F1, F2, F3) were measured for the vowel in each test word. A waveform editor was used to measure vowel durations from oscillographic displays by placing boundary markers at the first and last positive peaks in the periodic portions of the signals. The center frequencies of F1, F2, and F3 were estimated using linear predictive coding analysis by placing a 25.6 ms hamming window at the temporal midpoint of the vowel. F0 was calculated from the duration of three successive glottal periods at the vowel midpoint.

Statistical Analyses. The duration and frequency values of the English and German vowels were compared through three-way ANOVAs with group and gender as between-subjects measures and vowel as within-subject measure. Significant main effects revealed by ANOVAs were explored through Newman-Keuls post hoc tests ($\alpha < .05$). Two transformations of the formant frequency values in Hertz were carried out to normalize for gender differences. Transformation 1 converted the Hertz values into points along the three axes of Miller's (1989) Auditory–Perceptual Space. Transformation 2 converted the Hertz values into Barks and Bark-difference scores (Syrdal, 1985; Syrdal & Gopal, 1986). Both transformations of Hertz values were successful in removing gender main effects and interactions involving gender revealed by ANOVAs on untransformed Hertz values, and in retaining the vowel main effects and interactions that were the focus of the study. Results comparing the spectral properties of vowels will therefore be presented only for Bark-difference scores.[4] It has been shown that the Bark1–Bark0 (B1-B0) and Bark2–Bark1 (B2-B1) dimensions optimally separate front vowels (Syrdal, 1985; Syrdal & Gopal, 1986). These dimensions will be referred to here as *vowel height* (B1-B0) and *frontness–backness* (B2-B1).

Results

Vowels Produced by the Two German Groups. Before comparing English and German vowels in this way, it was first necessary to determine whether or not the difference in L2 experience between the two German groups had an effect on their production of German vowels. Previous studies have shown that L2 experience may affect the perception and production of L1 sounds (Anisfield, Anisfield, & Semogas, 1969; Anisfield & Gordon, 1971; Flege, 1987; Flege & Eefting, 1987a, 1987b; Flege & Hillenbrand, 1984; Schouten, 1977; Williams, 1979). Preliminary Group (2 levels) × Vowel (5 levels) ANOVAs revealed no significant differences ($p > .05$) between GA and GB for any of the acoustic variables examined (duration; F0, F1, F2, F3 in Hertz, or after transformations). This suggests that the 7-year difference in English language experience between the two German groups did not affect their production of German vowels. Therefore, the data for the two groups were pooled for the following analyses.

The mean durations for the German vowels produced by the German groups were as follows: /i/ (112 ms), /I/ (54 ms), /ɛ/ (76 ms), /ɛ:/ (163 ms), and /a/ (82 ms). A significant one-way ANOVA, $F(4, 76) = 214.84$, $p < .001$, was followed by post hoc tests, which showed that the durations for all vowels were significantly different except for /ɛ/ and /a/. The duration values clearly separate the long vowels /i, ɛ:/ from the short vowels /I, ɛ, a/.

The height dimension (B1-B0) and the front–back dimension (B2-B1) were examined in two one-way ANOVAs. The vowel effect was significant ($p < .001$) for both dimensions, $F(4, 76) = 281.91$ and $F(4, 76) = 387.49$, respectively. Post hoc tests showed that all of the vowels differed significantly along the two dimensions, except for /ɛ/ and /ɛ:/ along the height dimension.

It has been reported that regional differences exist in the use of vowel quality and duration to distinguish the vowels of German. To determine whether or not subjects in the two German groups differed in this respect, the 6 (of 20) subjects from northern Germany were compared to the 5 subjects from southern Germany. Differences between these two subgroups were tested in one-way ANOVAs examining variables that have been reported to vary with regional origin: spectral quality (F1, F2, F3, and Bark-difference scores for all five vowels; differences in Bark-difference scores for the vowel pairs /i/–/I/ and /ɛ/–/ɛ:/) and duration (absolute values for all vowels and duration ratios for the vowel pairs /i/–/I/ and /ɛ/–/ɛ:/). The comparisons did not reveal any major differences between the two regionally defined subgroups.[5] In particular, the subgroups did not differ for any of the spectral or temporal measures for /i/ and /I/ and for the /i/–/I/ contrast ($p > .10$). That is, the six northern and the five southern Germans made much the same use of duration and vowel quality in differentiating /i/ from /I/.

Vowels Produced by the Native English Group. Differences between the four English vowels were tested in a series of one-way ANOVAs. Main effects for B1-B0, $F(3, 27) = 174.61$, $p < .001$, and B2-B1, $F(3, 27) = 171.96$, $p < .001$, were obtained because all four vowels differed significantly along both the vowel height (B1-B0) and the front–back (B2-B1) dimension.

The mean durations for the four English vowels were as follows: /i/ (175 ms), /I/ (144 ms), /ɛ/ (181 ms), and /æ/ (238 ms). A main effect for duration was obtained, $F(3, 27) = 59.77$, $p < .001$, because the difference in duration was significant for all vowels, the exception being /i/ and /ɛ/. Unlike for the German vowels presented earlier, duration does not clearly separate vowels traditionally described as long (/i, æ/) from short (/I, ɛ/) vowels in English.

Acoustic Comparison of English /i, I, ɛ, æ/ and German Vowels. To help classify the four English vowels as new, similar, or identical from the standpoint of German, the English vowels /i, I, ɛ, æ/ as produced by the 10 native English speakers (EN) were compared to the German vowels /i, I, ɛ, ɛ:, a/ as produced by the 10 inexperienced Germans (GA).[6] Between-language comparisons were made only between those vowels that could reasonably be expected to be identified with one another by native German learners of English: the /i/s, /I/s, and /ɛ/s of English and German; English /ɛ/ and German /ɛ:/; and English /æ/ and German /ɛ, ɛ:, a/. Absolute vowel durations were not compared because the German tokens were disyllables and the English tokens monosyllables. Previous studies of English and German have shown that vowel duration is related to the number of syllables in a word (e.g., Harris & Umeda, 1974; Klatt, 1973; Meyer, 1904; Rietveld, 1975; Umeda, 1975). Instead, the duration ratios computed for pairs of English and of German vowels were compared.

English /i, I, ɛ/ and German /i, I, ɛ, ɛ:/. Figure 1 compares the distribution of the English vowels /i, I, ɛ/ as produced by EN to the German vowels /i, I, ɛ, ɛ:/ as produced by GA in the Bark-difference space. Separate one-way ANOVAs revealed significant differences in vowel height (B1-B0) between English /i/ and German /i/, $F(1, 18) = 14.72$, $p < .01$; between the /I/s of English and German, $F(1, 18) = 12.55$, $p < .01$; between the /ɛ/s of English and German, $F(1, 18) = 8.49$, $p < .01$; and between English /ɛ/ and German /ɛ:/, $F(1, 18) = 6.80$, $p < .05$. In all instances, the German vowels were higher in the acoustic vowel space than their English counterparts. In most instances, the English and German vowels did not differ significantly in frontness–backness (B2-B1). The one exception was English /ɛ/ and German /ɛ:/. German /ɛ:/ was more anterior in the acoustic vowel space than English /ɛ/, $F(1, 18) = 5.59$, $p < .05$. These results show partial spectral overlap for the English–German pairs /i/–/i/, /I/–/I/, and /ɛ/–/ɛ/. English /ɛ/ and German /ɛ:/ are also spectrally similar, but less so than the three pairs just mentioned.

As mentioned earlier, duration is used in both English and German to differentiate /i/ and /I/. However, the /i/–/I/ duration ratio is much larger in German (2.5) than in English (1.2), $F(1, 18) = 82.07$, $p < .001$. The /ɛ/ vowels of German and English are usually described as being short, but the present data indicate that this is phonetically appropriate only for German /ɛ/. With respect to duration, the long German /ɛ:/ is more similar to English /ɛ/ than is the short German /ɛ/. Although a direct comparison of absolute durations is problematic, it is of interest to note that the durations for German /ɛ:/ (163 ms) and for English /ɛ/ (181 ms) did not differ significantly, $F(1, 18) = 1.34$, $p > .05$.

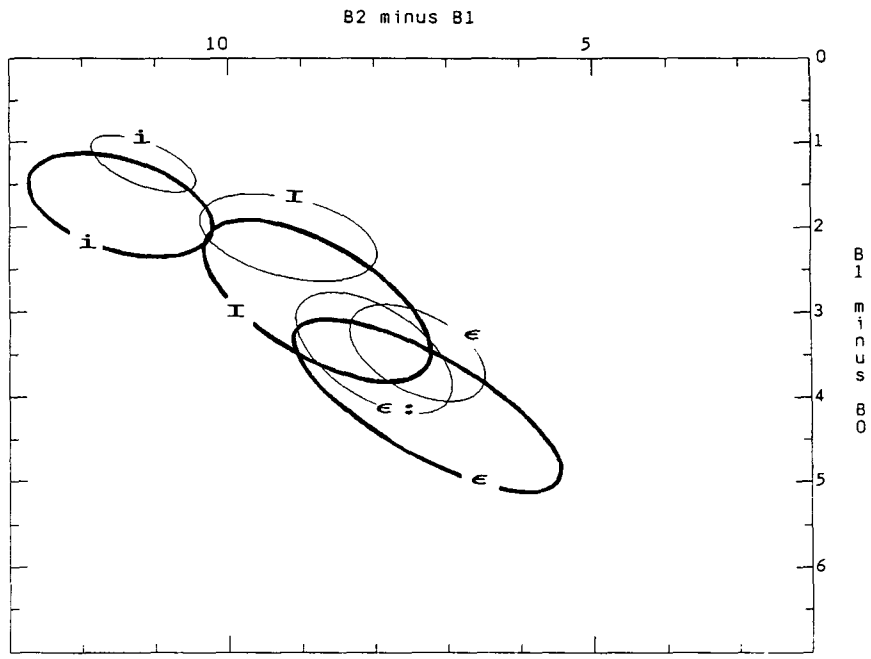

Figure 1. Range of the three English vowels /i, I, ϵ/ (as produced by monolingual native English speakers [bold lines]) and the German vowels /i, I, ϵ, ϵ:/ (as produced by native German speakers with relatively little English language experience [thin lines]) in the Bark-difference space. The ellipses enclose the mean Bark-difference values obtained for the 10 speakers in each group by representing 95% confidence levels based on the two principal components of variation for each vowel.

Relation of English /æ/ to German vowels. A spectral comparison of /æ/ to the nearest vowels of German suggests that it is a new vowel. Figure 2 shows the distribution of values for English /æ/ and those for German /ϵ, ϵ:, a/. The plot clearly shows little overlap between /æ/ and any of the German vowels for the German subjects in this study. Because English /æ/ has no clear counterpart in German, no formal statistical analyses were carried out, but it is obvious that English /æ/ differs significantly from any German vowel it might be compared to. On the basis of mean absolute vowel durations, it seems that English /æ/ is longer (238 ms) than German /ϵ/ (76 ms), /a/ (82 ms), and /ϵ:/ (163 ms).

Conclusion

The acoustic comparisons indicate that English /i/ and /I/ should be classified as similar for native Germans because they resemble German /i/ and /I/ spectrally and temporally but are *not* identical to them. English /ϵ/ is similar to both German /ϵ/ and /ϵ:/. English /ϵ/ is more similar to German /ϵ/ than /ϵ:/ spectrally, but

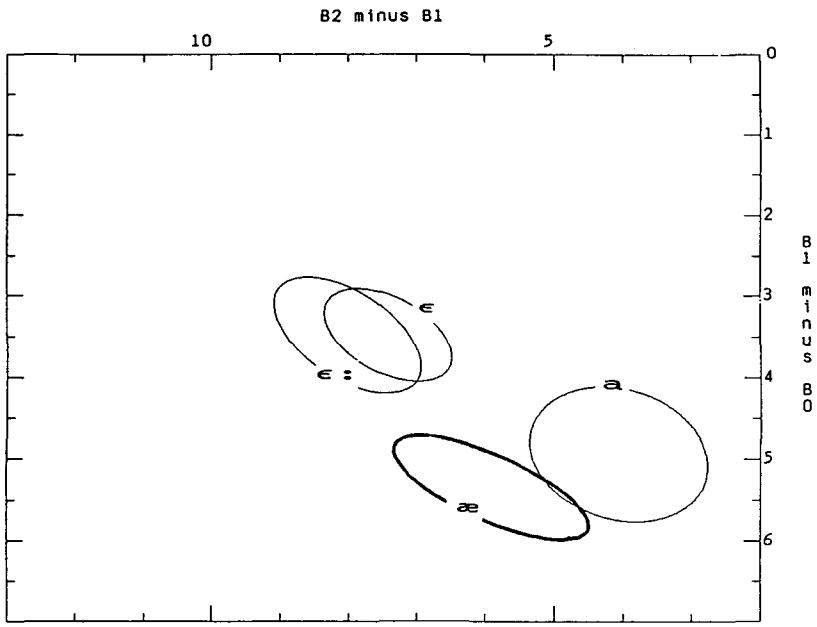

Figure 2. Range of the English vowel /æ/ (as produced by monolingual native English speakers [bold line]) and the German vowels /ɛ, ɛ:, a/ (as produced by native German speakers with relatively little English language experience [thin lines]) in the Bark-difference space.

temporally it resembles German /ɛ:/ to a greater extent than German /ɛ/. It is uncertain how spectral and temporal properties of vowels should be weighted in a cross-language comparison of vowels. The English vowel /æ/, on the other hand, is a new sound for native Germans.

These classifications of English vowels from the standpoint of German were made on the acoustic–phonetic, not a phonological level because the present study evaluates a model of L2 speech learning that states that interlingual identification occurs on a phonetic level, and because this model makes predictions about measurable phonetic properties of L2 learners' speech. (For detailed expositions of the model, especially of crucial concepts such as *phonetic similarity* and *equivalence classification*, see Flege, 1988, 1991.) It should be noted that the present acoustically based classifications of English /i, I, ɛ/ as similar and English /æ/ as new for L1 Germans are fully consistent with the results of a perceptual study examining interlingual identifications of English /i, I, ɛ, æ/ by near-monolingual German listeners (Bohn & Flege, 1990).

Based on these classifications, the model being evaluated leads to several testable predictions. First, amount of English language experience should not substantially affect L1 German speakers' productions of the similar English vowels /i, I, ɛ/. That is, subjects in both the GB and the GA group will differ from EN speakers in produc-

ing English /i, I, ɛ/. The prediction for the new English vowel /æ/, on the other hand, is that English L2 experience will enable at least some L1 German speakers to achieve the English acoustic norm. It is predicted that speakers in GB will differ less from EN speakers than speakers in GA, and that GB speakers may in fact match the EN speakers in their /æ/ productions. These predictions were tested through acoustic measurements (Experiment 2) and in a perception experiment (Experiment 3).

EXPERIMENT 2: ACOUSTIC COMPARISON OF ENGLISH /i, I, ɛ, æ/ AS PRODUCED BY NATIVE ENGLISH AND NATIVE GERMAN SPEAKERS

The aim of Experiment 2 was to compare native English and native German speakers for acoustic properties of their productions of English /i, I, ɛ, æ/. As in Experiment 1, the analyses examined determinants of vowel identity (spectral properties at the vowel midpoint, duration) and relational properties of vowel pairs.

Method

The same L1 German and L1 English speakers as in Experiment 1 participated. The subjects were all recruited from the same university community (see Experiment 1) so that the native Germans' English productions would be compared to L2 "norms" that may be assumed to represent L2 "targets" for the native Germans. The speech materials recorded from the native English (EN) group for Experiment 1 were also used for Experiment 2. Recordings of English *bVt* words (V = /i, I, ɛ, æ/) from the experienced (GB) and the inexperienced (GA) native German speakers were made using the same procedures as for EN in Experiment 1.[7] Acoustic measurements and statistical analyses followed the same procedures as for Experiment 1.

Results

The speech production data were first analyzed in separate Group (3 levels) × Vowel (4 levels) ANOVAs with vowel duration and Bark-difference scores (B1-B0 and B2-B1) as dependent variables. The ANOVAs revealed significant Group × Vowel interactions for vowel duration, $F(6, 81) = 4.03$, $p < .001$, and for the height dimension B1-B0, $F(6, 81) = 3.34$, $p < .01$. There was a main effect of vowel for the front–back dimension B2-B1, $F(3, 81) = 459.09$, $p < .001$. However, the Group × Vowel interaction for B2-B1 was nonsignificant, $F(6, 81) = 2.21$, $p > .05$, as was the group effect for B2-B1, $F(2, 27) = 0.22$, $p > .05$. These results indicate that the groups differed significantly in the production of individual vowels in terms of vowel duration and vowel height (B1-B0) but not in terms of frontness–backness (B2-B1).

The Group × Vowel interactions were explored through one-way ANOVAs testing the effect of group (EN, GB, GA) on the durations and the B1-B0 scores of individual vowels. Because neither the Group factor nor the Group × Vowel interaction was significant for B2-B1, these scores were not further analyzed in the comparisons of English vowels. Separate one-way ANOVAs for each German group were

also carried out on the Bark-difference scores and duration ratios to determine the effect of language (German vs. English) on the production of /i, I, ɛ/ in German versus English.

Acoustic Comparison of English /i/ and /I/ Produced by Native German and Native English Speakers.

Comparison of spectral properties. The top panel of Figure 3 shows the range of English /i/ and /I/ as produced by EN and GB; the bottom panel juxtaposes the EN values and those of GA. The English /i/s and /I/s produced by both German groups were higher in the acoustic space than those of EN. The effect of group on the B1-B0 scores was significant for /i/, $F(2, 27) = 5.23$, $p < .05$, because the /i/ produced by EN was significantly lower (with a mean B1-B0 score of 1.72) than the /i/s produced by GB and GA (mean B1-B0 scores of 1.29 for both groups). The group effect on the B1-B0 scores for /I/ was nonsignificant, $F(2, 27) = 3.22$, $p > .05$.

Additional one-way ANOVAs were carried out to determine whether or not the two German groups produced /i/ and /I/ differently in English and German. For both German groups, the effect of language on B1-B0 and B2-B1 was nonsignificant for /i/, and the language effect was also nonsignificant for B2-B1 for /I/ ($p > .05$). However, the language effect was significant for B1-B0 of /I/ produced by GB, $F(1, 18) = 5.92$, $p < .05$. Their English /I/ had a larger B1-B0 value (2.33) than their German /I/ (1.98), which means that they produced a lower and thus more English-like /I/ in English than in German. Although all other German–English differences in B1-B0 scores for /i/ and /I/ were nonsignificant, it is worth noting that they all pointed in the same direction: phonetic approximation by means of a lowering of /i/ and /I/ in English compared to German.

Comparison of vowel durations. Figure 4 shows the mean durations of English /i/ and /I/ as produced by the subjects in EN, GB, and GA. Juxtaposed to the English values are the mean durations of German /i/ and /I/ as produced by GB and GA (the two rightmost columns). As reported earlier, German /i/ was longer than German /I/, and the speakers in EN also made English /i/ longer than English /I/. The duration difference between /i/ and /I/, however, was smaller in English than in German. Figure 4 shows that the Germans produced a larger English /i/–/I/ duration contrast than the EN speakers. This seems to be due mainly to the relatively short English /I/ durations produced by the native Germans.

Differences among the speakers in GA, GB, and EN for the durations and duration ratios of English /i/ and /I/ were examined in one-way ANOVAs. The group means for the duration of English /i/ were 175 ms for EN, 155 ms for GB, and 175 ms for GA. These differences were nonsignificant, $F(2, 27) = 0.98$, $p > .05$. However, the groups differed significantly for the duration of English /I/, $F(2, 27) = 4.90$, $p < .05$, because the speakers in EN produced significantly longer /I/ durations ($M = 144$ ms) than the speakers in GB ($M = 91$ ms) and GA ($M = 102$ ms), who did not differ significantly from one another.

The comparison of the duration ratios of English /i/ to /I/ revealed a significant group effect, $F(2, 27) = 8.12$, $p < .01$, because the ratio for the speakers in EN (1.22) was significantly smaller than the ratio for speakers in GB (1.70) and GA (1.72), who

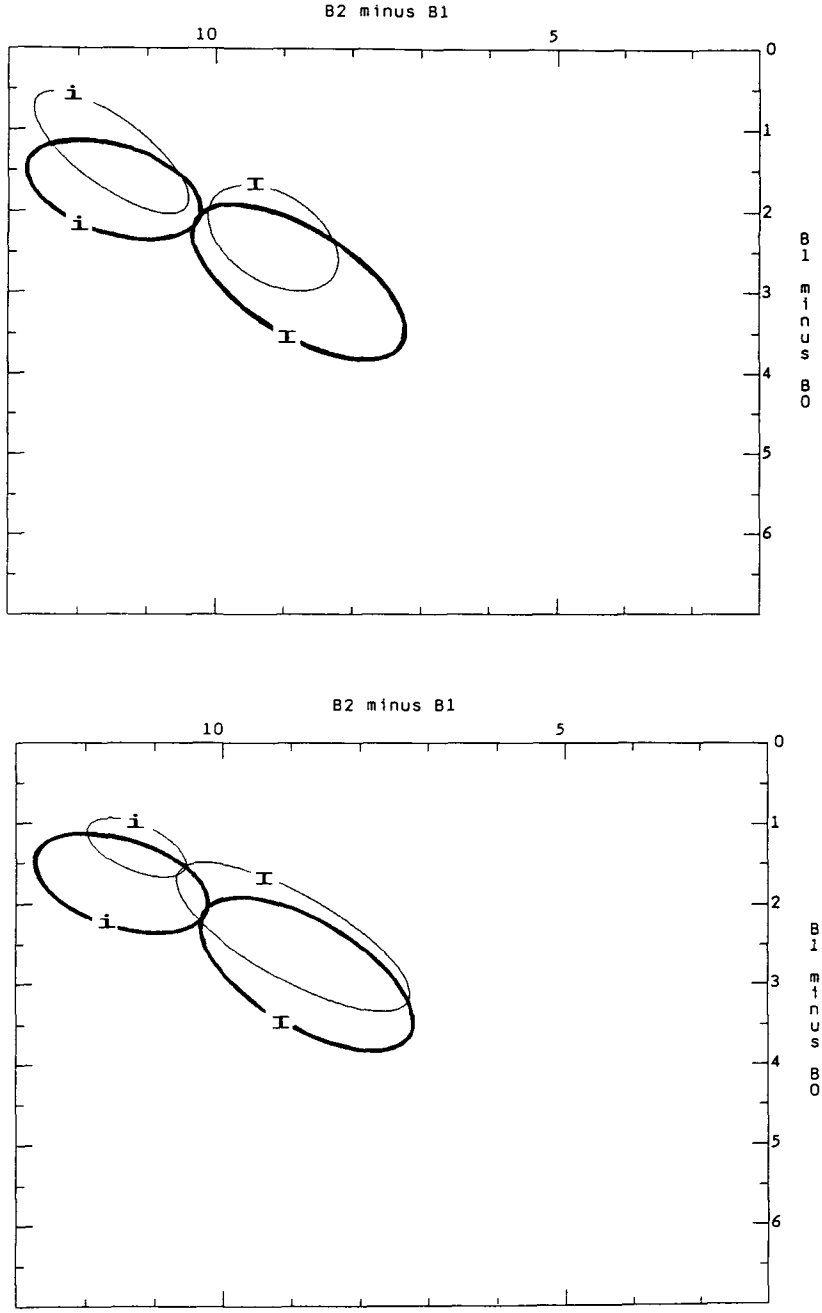

Figure 3. Range of the English vowels /i/ and /I/ as produced by native English speakers and experienced German speakers of English (top panel) and as produced by native English speakers and inexperienced German speakers of English (bottom panel) in the Bark-difference space. (Bold lines: native English speakers; thin lines: native German speakers.)

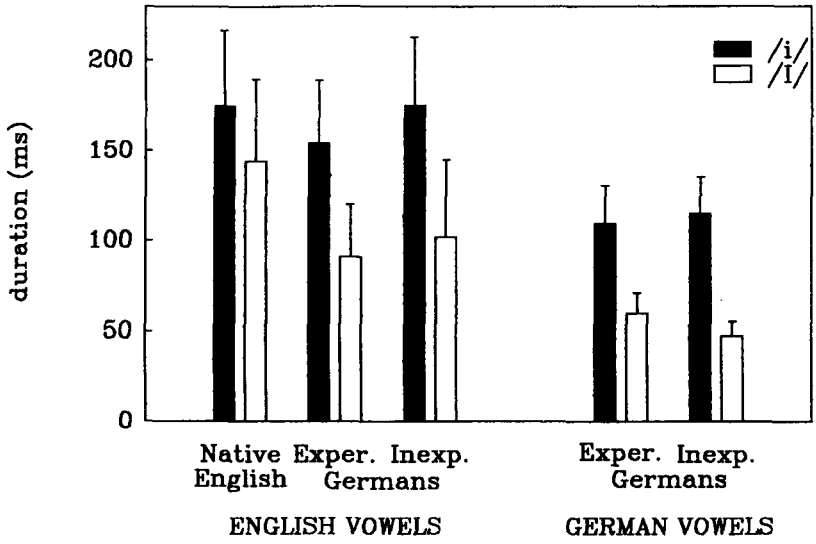

Figure 4. Mean durations of English /i/ and /I/ as produced by the native English and the two German groups differing in English language experience (leftmost columns) and mean durations of German /i/ and /I/ as produced by the native German groups (rightmost columns). Error bars indicate 1 standard deviation.

did not differ significantly from one another. The comparisons for the German groups across languages revealed that the /i/ to /I/ duration ratio for GA was significantly smaller in English (1.72) than in German (2.45), $F(1, 18) = 8.67, p < .01$. The speakers in GB did not differ significantly for the duration ratios they produced for /i/ and /I/ in English (1.70) versus German (1.83), $F(1, 18) = 0.35, p > .05$.

Acoustic Comparisons of English /ɛ/ and /æ/ Produced by Native German and Native English Speakers.

Comparison of spectral properties. Figure 5 shows the range of English /ɛ/ and /æ/ as produced by speakers in EN (top panel), GB (center panel), and GA (bottom panel) in the acoustic vowel space. The plot shows a fairly clear separation of /ɛ/ and /æ/ productions for EN and GB and an almost complete overlap for GA, indicating that the experienced, but not the inexperienced, Germans produced an English-like /ɛ/–/æ/ contrast.

To determine between-group differences in the separation of the /ɛ/ and /æ/ categories, the difference in vowel height (B1-B0 for /æ/ minus B1-B0 for /ɛ/) was calculated for each subject and submitted to a one-way ANOVA. A significant group effect, $F(2, 27) = 10.28, p < .001$, was obtained because the speakers in GA produced a smaller vowel height difference between /ɛ/ and /æ/ ($M = 0.22$) than the speakers in EN and GB, who did not differ significantly from one another (mean difference in B1-B0 scores for EN: 1.25, for GB: 1.24).

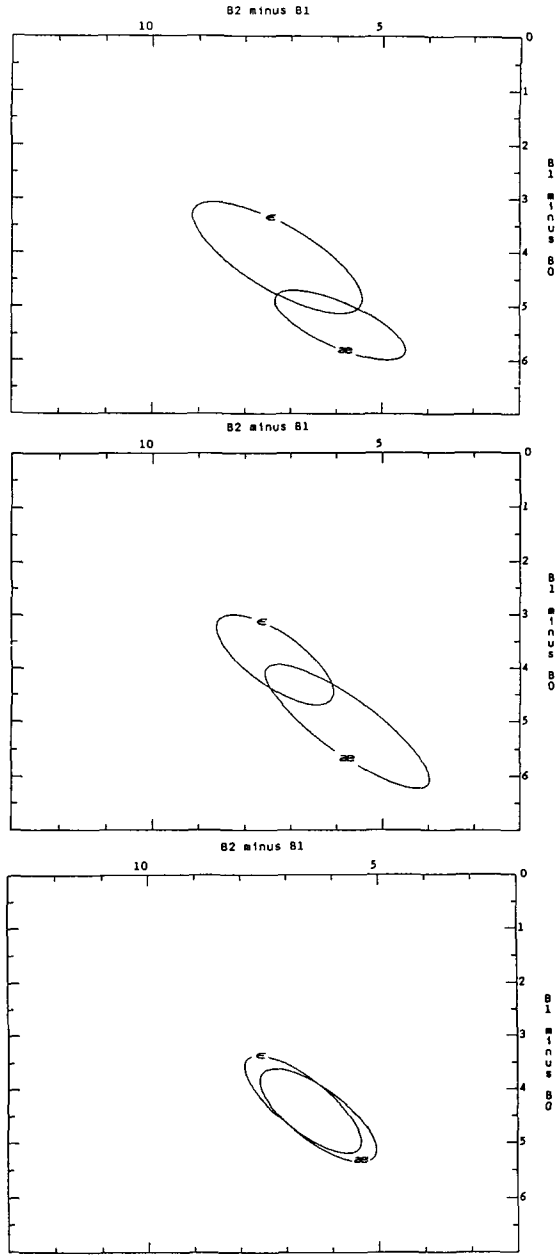

Figure 5. Range of the English vowels /ɛ/ and /æ/ as produced by the native English group (top panel), the experienced German group (center panel), and the inexperienced German group (bottom panel) in the Bark-difference space.

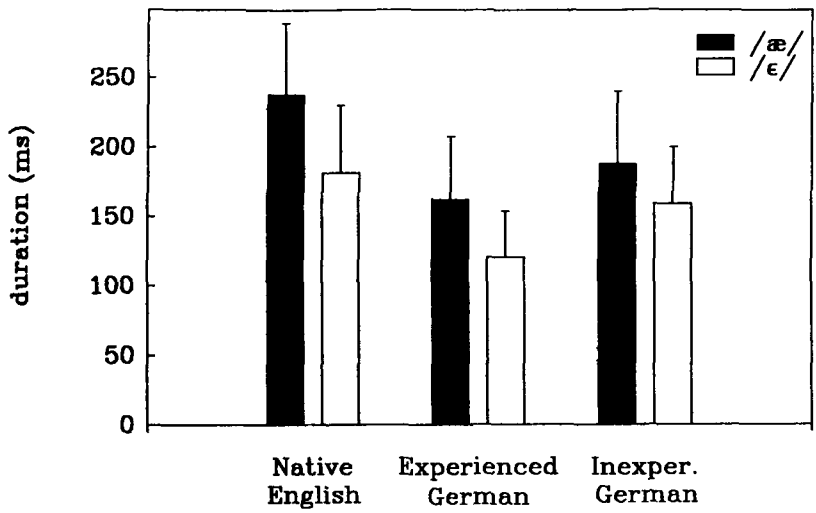

Figure 6. Mean durations of English /ɛ/ and /æ/ as produced by the native English and the two German groups differing in English language experience. Error bars indicate 1 standard deviation.

Focusing first on English /æ/, which was shown to be a new vowel for GA and GB, a significant group effect was obtained for the height dimension, $F(2, 27) = 6.65$, $p < .01$, because the /æ/ produced by GA was significantly higher (with a mean B1-B0 score of 4.50) than the /æ/ produced by EN and GB. The B1-B0 scores for GB ($M = 5.07$) did not differ significantly from those for EN ($M = 5.35$). The comparison of the Bark-difference scores for English /ɛ/ revealed no significant group effects, which indicates that both German groups produced a spectrally native-like English /ɛ/.

The comparison of English to German vowels indicated that German has two vowels that are similar to English /ɛ/, namely, /ɛ/ and /ɛ:/. Separate one-way ANOVAs comparing the two native German groups' English /ɛ/ to their German /ɛ/ and /ɛ:/ revealed significant differences only for GA. Their English /ɛ/ was both lower and more anterior than their German /ɛ/ and their German /ɛ:/.[8]

Comparison of vowel durations. Figure 6 shows the mean durations of English /ɛ/ and /æ/ as produced by the speakers in EN, GB, and GA. The group effect for the duration of /ɛ/ was significant, $F(2, 27) = 5.58$, $p < .01$, because the subjects in GB produced shorter /ɛ/s ($M = 120$ ms) than those in EN ($M = 181$ ms) and in GA ($M = 159$ ms), who did not differ significantly from one another. The group effect for the duration of /æ/ was significant, $F(2, 27) = 5.95$, $p < .01$, because the EN speakers produced longer /æ/s ($M = 238$ ms) than the speakers in GB ($M = 162$ ms) and GA ($M = 188$ ms), who did not differ significantly from one another.

The comparison of the duration ratios of English /æ/ to /ɛ/ revealed a barely significant group effect, $F(2, 27) = 3.35$, $p = .0496$. Post hoc tests did not reach significance for any between-group comparison. Pairwise comparisons of the three

groups in one-way ANOVAs studying the effect of group (2 levels) on /æ/–/ɛ/ duration ratios indicated that the ratio for GA ($M = 1.18$) was significantly smaller than the ratio for EN ($M = 1.31$), $F(1, 18) = 5.27$, $p < .05$, and for GB ($M = 1.35$), $F(1, 18) = 5.84$, $p < .05$. The /æ/–/ɛ/ duration ratios for EN and GB did not differ significantly from one another, $F(1, 18) = 0.18$, $p > .05$.

Conclusion

The results for /i/ and /I/ are consistent with the hypothesis that amount of L2 experience does not affect the production of similar L2 vowels. No differences were found between GB and GA with regard to vowel spectrum and absolute and relative duration of their English /i/ and /I/ productions. If vowel duration and spectrum are considered together, the native Germans did not reach the acoustic norms of English as defined by EN for either /i/ or /I/. Their English /i/ had an appropriate duration but was higher than the /i/ produced by EN. The Germans' English /I/ did not differ spectrally from the /I/ produced by EN, but it was inappropriately short compared to the /I/ produced by EN. The Germans also produced a larger duration contrast than the speakers in EN. However, both German groups showed some evidence of approximating the acoustic norms of English. The speakers in GB produced a lower (i.e., more English-like) /I/ in English than in German. The speakers in GA produced a smaller /i/–/I/ duration contrast in English than in German.

The results for /æ/ are largely consistent with the hypothesis that extended L2 experience will enable adults to produce a new vowel in a nativelike fashion. The experienced, but not the inexperienced, Germans produced an /æ/ vowel that did not differ significantly from the native English speakers' /æ/ in terms of spectral quality. However, the GB speakers produced significantly shorter /æ/s than the speakers in EN. While the absolute /æ/ durations do not support the hypothesis, the duration ratios of /æ/ to /ɛ/ indicate that the experienced, but not the inexperienced, Germans had established an English-like duration contrast between the two spectrally neighboring vowels.

The results for English /ɛ/ are not consistent with the hypothesis that L2 experience does not affect the production of similar vowels. The English /ɛ/ of both German groups did not differ spectrally from the native English speakers' /ɛ/, although both German /ɛ/ and /ɛ:/ were found to be higher in the acoustic vowel space than English /ɛ/. Another unexpected result was that only the *in*experienced Germans produced /ɛ/ durations that were English-like, whereas the experienced Germans produced shorter /ɛ/s than both the native English and the inexperienced German group.

The finding that inexperienced L2 speakers were more successful at producing a similar L2 vowel than experienced ones seems paradoxical. It appears in a different light, however, if the results for /æ/ are considered. As shown in Figure 5, the GA speakers did not differentiate between English /ɛ/ and /æ/, which suggests that they used only one vowel category where the speakers in EN and GB used two. It seems likely that the significant difference for GA between their native German /ɛ/

and the similar English /ɛ/ is due to a merger of the acoustic properties of /ɛ/ and the new vowel /æ/, which resulted in an English /ɛ/ that was lower in the vowel space than their native /ɛ/ but not as low as English /æ/.

EXPERIMENT 3: INTELLIGIBILITY FOR NATIVE ENGLISH LISTENERS

The predictions presented initially were further tested in an experiment examining how well native English speakers could identify the English vowels in *beat*, *bit*, *bet*, and *bat* as produced by the subjects in GA, GB, and EN. It was expected that *beat*, *bit*, and *bet* spoken by the GA and GB would be identified correctly, as English /i, I, ɛ/ were shown to be similar to German vowels. No differences in intelligibility were expected to exist between GA and GB for these vowels. On the other hand, the intended *bat* productions of GB were expected to be more intelligible (i.e., identified correctly in a larger percentage of instances) than those of GA, because /æ/ is a new vowel.

Method

Three adult monolingual native speakers of American English served as paid listeners. They heard the English words *beat*, *bit*, *bet*, and *bat* as produced by the 10 subjects each in GA, GB, and EN in separate blocks. The *bVt* words had been normalized for peak vowel intensity. The 20 tokens (4 words × 5 repetitions) available for each speaker were randomly presented five times each. The listeners used a response box to indicate which word they heard: *beat*, *bit*, *bait*, *bet*, *bat*, *but*, or *bot(tle)*. Six speakers were presented on each of 5 days. The order in which the speakers were presented to the listeners was counterbalanced across listeners, with the provision that an equal number of speakers from each group be presented on each day.

This procedure yielded a total of 3,000 responses (3 groups × 10 speakers × 20 tokens × 5 randomizations). The rate at which the vowels were correctly identified was tabulated in separate response matrices for each speaker. Each cell in the matrix represented the average for the three native English listeners. The percentage scores were analyzed in nonparametric statistical tests (Kruskal-Wallis) to determine between-group differences in the rate of correct identifications of intended *beat*, *bit*, *bet*, and *bat* productions.

Results

Figure 7 shows the percentage of correct identification for the four English vowels. As expected, *beat* and *bit* as produced by GA and GB were as intelligible as the EN speakers' *beat* and *bit*. There were no significant differences ($\alpha < .05$) between the three groups for *beat* (EN: 100%, GB: 100%, GA: 99.9%) or for *bit* (EN: 100%, GB: 99.6%, GA: 98.7%). However, the native Germans' /ɛ/s and /æ/s were less intelligible than those of the native English speakers.

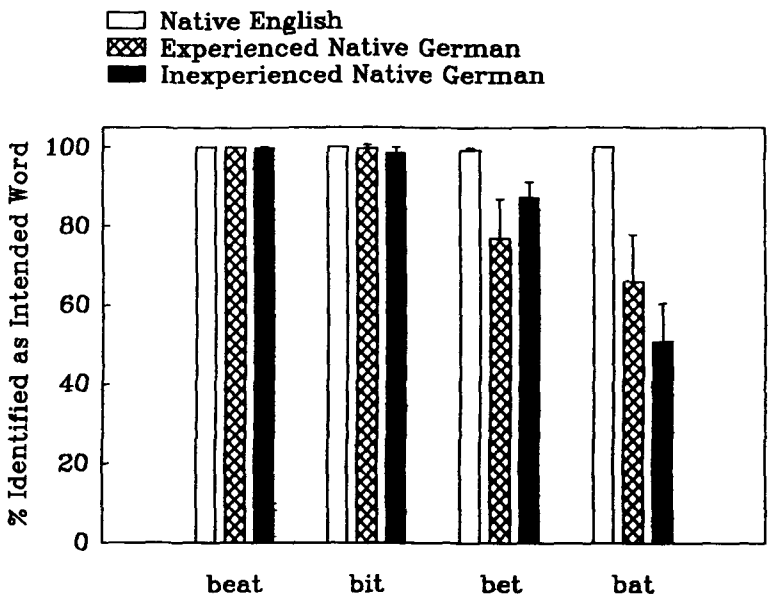

Figure 7. Percentage of correct identification of English *beat*, *bit*, *bet*, and *bat* as produced by a group of native English speakers and two groups of native German speakers differing in English language experience. Error bars indicate 1 standard deviation.

There was a significant difference among the three groups for *bet*: EN (99%), GB (76.9%), GA (87.5%), $H = 11.89$, $p < .01$. Pairwise comparisons revealed that the EN group differed from both German groups, EN versus GA: $H = 7.79$, $p < .01$; EN versus GB: $H = 11.80$, $p < .001$, but the two German groups did not differ significantly from one another, $H = 0.09$, $p > .05$. There was a significant difference in the rate of correct identifications for intended *bat* tokens among EN (100%), GB (65.9%), and GA (50.9%), $H = 14.77$, $p < .001$. Pairwise comparisons revealed that EN differed from both German groups, EN versus GB: $H = 9.64$, $p < .01$; EN versus GA: $H = 13.87$, $p < .001$. Contrary to the prediction, the /æ/ produced by GB was not significantly more intelligible than /æ/ produced by GA, $H = 1.05$, $p > .05$.

The relatively poor intelligibility of English /ɛ/ for the German subjects is surprising in one sense because /ɛ/—like the English vowels /i/ and /ɪ/, which were identified quite well—was classified as a similar vowel. The results are understandable, however, with reference to the acoustic comparison of German /ɛ/ to English /ɛ/ (earlier). Recall that German /ɛ/ was shorter in duration and higher in the acoustic vowel space than English /ɛ/. In a recent study of interlingual vowel perception, Bohn and Flege (1990) found that near-monolingual German listeners identified the vowel in the English word *bet* as produced by two native English speakers in most cases using a German word with /ɛ/ (*Bett*). They less frequently used *bäht*, which has [ɛː], or *Bert*, which may have [ɛː] or [æ(ː)] in the dialect of the listeners in

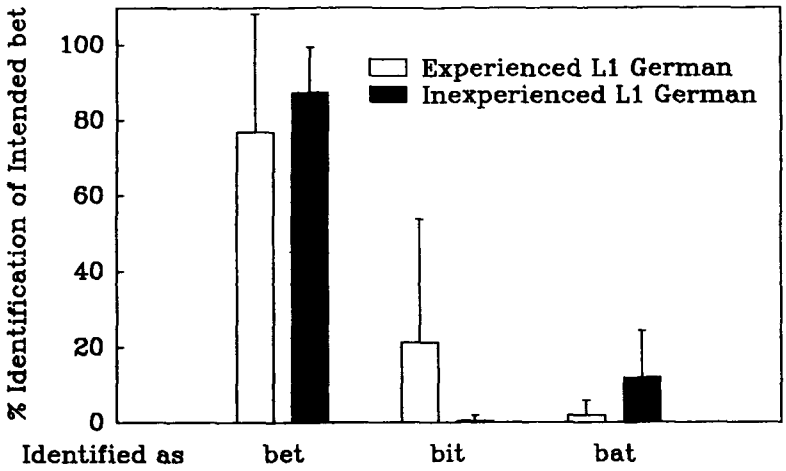

Figure 8. Percentage of identification of the native German speakers' intended *bet* as *bet*, *bit*, or *bat*. Error bars indicate 1 standard deviation.

the Bohn and Flege (1990) study. Thus, diminished intelligibility for English *bet* is just what one would expect if the German subjects used their German /ɛ/ to produce this word.

This conclusion may apply better to GB than to GA. Figure 8 shows in more detail the categorizations of the L1 German speakers' intended *bet* tokens. There were significantly more *bit*-for-*bet* misidentifications for GB than for GA, $H = 3.87$, $p < .05$. (This result was entirely due to the productions of 4 of the 10 GB subjects.) Conversely, there were significantly more *bat*-for-*bet* misidentifications for GA than GB, $H = 7.27$, $p < .01$. It appears that subjects in GB were more likely than subjects in GA to use their higher and shorter German /ɛ/ in English *bet*, and that more subjects in GA than in GB may have used a vowel in *bet* that differed from their native /ɛ/.[9]

The results for *bat* do not support the prediction that extended L2 exposure will result in intelligible productions of a new vowel. Only six subjects from the two German groups produced *bat* tokens whose percentage of correct identification exceeded 90%. Of these, four were experienced learners. The remaining Germans' *bat* tokens, when misidentified, were heard as *bet*. This is consistent with previous observations of sound substitutions by German learners of English (Arndt & Careless, 1978; Barry, 1977, 1981; Oakeshott-Taylor, 1976; Weiher, 1975; Wode, 1978, 1981).

COMPARISON OF INTELLIGIBILITY AND ACOUSTIC RESULTS

The results of the intelligibility judgments and of the acoustic comparisons are consistent with respect to the L1 Germans' productions of English /i/ and /ɪ/. However, there are discrepancies between the perceptual and acoustic assessments of English /ɛ/ and /æ/ as produced by the German groups. Although previous L2 speech

studies have revealed both theoretical and practical problems in comparing categorical perceptual judgments to acoustic measurements (e.g., Barry, 1977), an attempt will be made to relate and evaluate the discrepant findings of the two experiments for /ɛ/ and for /æ/.

The perceptual results suggested that both German groups produced a less intelligible English /ɛ/ than EN. The acoustic comparisons revealed that both German groups did not differ significantly from EN in terms of spectral properties of /ɛ/. Significant duration differences existed only between GB on the one hand and GA and EN on the other. These findings make one wonder why the GA speakers' /ɛ/ was less intelligible than that of the EN speakers, even though their /ɛ/ did not differ significantly for the duration or spectral properties measured. It is of course possible that the intelligibility judgments were influenced by acoustic properties of the *bet* tokens that were not acoustically analyzed in the present study, such as formant movements. However, it appears likely that the intelligibility results for GA were related to the lack of acoustic differentiation between /ɛ/ and /æ/ because virtually all misidentifications of intended *bet* by GA speakers were heard as *bat* (and all misidentifications of intended *bat* as *bet*). The responses of the judges in the perception experiment, who also heard perfectly intelligible /ɛ/s and /æ/s from the speakers in EN and GB, may well have been influenced by expectations to hear both /ɛ/s and /æ/s from each individual speaker.

The most important discrepancy between the two experiments concerns the vowel /æ/. Both German groups produced shorter /æ/s than the EN group, but only the inexperienced Germans differed from the EN speakers in terms of vowel height, the difference in vowel height between /æ/ and /ɛ/, and in terms of the /æ/–/ɛ/ duration ratio. These acoustic results, which indicated that the GB speakers produced an /æ/ that was much like the /æ/ of the EN speakers, were largely consistent with the initially stated hypothesis that L2 experience would lead to the establishment of a new vowel category. However, the perception experiment indicated that both German groups produced less intelligible /æ/s than EN. There was only a nonsignificant trend for GB speakers to produce a more intelligible /æ/ than the GA speakers.

These findings suggest that the intelligibility results for the German groups may have been related to their relatively short /æ/ durations. Correlational analyses comparing percentages of /æ/ identifications to /æ/ durations indicated a moderate relationship between the criterion variable and /æ/ durations for GA, $r = .59$, $p < .05$. The correlation coefficient for GB was nonsignificant, $r = .42$, $p > .05$. Additional correlational analyses were carried out to determine whether or not /æ/ intelligibility was related to vowel height for /æ/ and to the difference in vowel height between /æ/ and /ɛ/. For GB, high correlations were obtained between the criterion variable and (a) vowel height, $r = .80$, $p < .01$, and (b) the difference in vowel height between /æ/ and /ɛ/, $r = .83$, $p < .01$. The correlations for GA were lower and significant only for the criterion variable and the height difference between /æ/ and /ɛ/, $r = .62$, $p < .05$, but not for percentage of correct /æ/ identifications and the height of /æ/, $r = .30$, $p > .05$.

These comparisons suggest that the listeners may have used different phonetic

criteria in categorizing intended *bat* tokens of GB as opposed to GA speakers. Because the listeners used only two response alternatives for categorizing intended *bat* tokens (viz., *bet*, *bat*), the present results may be compared to those of an earlier study (Bohn & Flege, 1990) in which members from a synthetic *bet–bat* continuum varying factorially in vowel spectrum and duration were presented for identification as *bet* or *bat*. Bohn and Flege found that the responses of native English listeners were based almost exclusively on the spectral properties of the stimuli, whereas duration had little influence on the responses. A possible conclusion for the present listening experiment is that the native English listeners used the primary phonetic criterion for perceptual differentiation between /æ/ and /ɛ/ (i.e., vowel spectrum) to categorize the intended *bat* tokens of GB. The intended *bat* tokens of GA, however, which were acoustically much like their intended *bet* tokens, were apparently categorized not using only the primary acoustic cue of vowel spectrum. In addition, the listeners also used vowel duration, a cue not normally used by native English listeners to differentiate /æ/ from /ɛ/ in perception (see also Lieberman & Kubaska, 1979).

GENERAL DISCUSSION

The primary purpose of this study was to examine whether or not the production of L2 sounds is influenced by amount of L2 exposure and by sound correspondences between the L1 and L2. This was done through acoustic and perceptual assessments of English vowel productions by two groups of native German speakers differing primarily in English language experience. The L2 speakers of English, and a monolingual native English control group, produced the English vowels /i, I, ɛ/, which were shown to be similar, but not identical, to German /i/, /I/, and /ɛ/ or /ɛ:/, and they produced the English vowel /æ/, which was shown to be acoustically quite unlike any German vowel.

The hypothesis tested in this study was derived from Flege's (1987, 1988, 1991) speech learning model. It predicts that adult learners will eventually produce new L2 sounds (like English /æ/ for L1 Germans) authentically because extended L2 experience leads to the establishment of phonetic categories for L2 sounds that have no counterpart in the L1. Similar sounds (like English /i, I, ɛ/ for L1 Germans), however, will remain foreign-accented even after lengthy L2 exposure. This is hypothesized to be so because category formation for similar sounds in adults is blocked by equivalence classification, a mechanism that causes acoustically similar sounds in the L1 and L2 to be treated as belonging to the same phonetic category.

The predictions of the model were tested in two experiments that examined specific aspects of authenticity in L2 speakers' vowel productions. One experiment examined authenticity in terms of acoustic determinants of vowel identity (spectral properties at the vowel midpoint, vowel duration) and vowel contrasts (spectral differences and duration ratios of vowel pairs). Although results based on these measures present important evidence bearing on the models' predictions, it should be noted that acoustic vowel properties that were not examined (e.g., formant dynamics) may also contribute to authenticity in as yet unknown ways. The percep-

tual experiment examined authenticity in terms of intelligibility, which is a fairly coarse-grained perceptual measure compared to measures calling for subphonemic judgments (e.g., acceptability ratings). In theory, it would have been preferable to examine authenticity more directly in terms of acceptability. This was considered inappropriate, however, because listeners would have been required to provide within-category judgments of tokens from a corpus containing misidentifiable tokens. In interpreting the intelligibility results, it should be noted that intelligibility is a necessary, not a sufficient, prerequisite for L2 productions to be considered authentic. That is, any authentic production should be intelligible, but intelligible productions may or may not be foreign accented.

As predicted, the perceptual and acoustic results showed that the production of English /i/ and /I/ was not substantially affected by L2 experience. The two German groups largely retained properties of German /i/ and /I/ in their English /i/ and /I/ productions. That is, they produced a higher English /i/ than the native English group, and the duration contrast between English /i/ and /I/ was larger for the native German than the EN speakers. However, both German groups produced /i/s and /I/s in English that were closer to the English acoustic norm than their German /i/s and /I/s, but phonetic approximation of the EN speakers' values was generally too small in magnitude to be significant and/or to result in acoustic values that would match the English acoustic norm.

The results for /i/ and /I/ underscore the generally held belief that these English vowels are unproblematic for L1 Germans. However, the results also suggest that phonetic learning for similar sounds may be largely restricted to the early stages of L2 contact. This finding seems to be at variance with a result in the Flege (1987) study, which examined native English speakers' productions of French sounds, including the similar /u/ vowel. Flege reported that only the most experienced English speakers of French approximated the French acoustic norm for /u/, which is more posterior in the acoustic vowel space than English /u/. As will be argued later (in the discussion of the findings for /ɛ/), the result for /u/ was probably obtained because the French /u/ of those of Flege's subjects who had relatively little French language experience was "deflected" to a more anterior position than their English /u/ by the new French vowel /y/, for which they had not yet established a phonetic category. Taken together, Flege's (1987) findings and those of the present study suggest that phonetic learning for a similar sound takes place early during L2 exposure and does not progress much along with L2 experience if the similar sound is not in the acoustic neighborhood of a new sound (as for English /i/ and /I/ in the present study). However, if a similar sound (like French /u/) is in the acoustic neighborhood of a new sound (like French /y/), experienced speakers may approximate the acoustic norm for the similar sound once they have established a phonetic category for the neighboring new sound. This hypothesis, which needs further testing, receives some support from additional results reported in Flege (1987) and from the present results for /ɛ/ (see later).

The present finding for /i/ and /I/ that continued L2 exposure does not lead to continued phonetic learning beyond a certain point for similar vowels could be interpreted as being due to the workings of the general principle of least effort. The

perceptual results indicated that the speakers in both GA and GB produced /i/s and /I/s in English that were as highly intelligible as the EN speakers' /i/s and /I/s. One might therefore suspect that phonetic learning for /i/ and /I/ stops early during L2 exposure because further approximation toward the English norms would not further improve intelligibility. The present results for /ɛ/, however, show that this view is incorrect. The speakers in GB approximated, but did not reach, the English acoustic norm for this similar vowel although continued learning would have improved the intelligibility of their /ɛ/s. We may conclude, therefore, that Flege's Speech Learning Model is correct in stating that equivalence classification blocks phonetic learning for similar vowels beyond a certain point, and that it is not communicative adequacy that determines at what point phonetic learning stops.

The results for /ɛ/ were mixed with respect to the hypothesis tested. As predicted for similar vowels, the two German groups differing in English L2 experience did not differ in intelligibility for English /ɛ/. However, contrary to the prediction, the speakers in GA produced an English /ɛ/ that matched the EN speakers' /ɛ/ for spectral and duration properties, whereas the GB speakers matched the EN speakers only for vowel spectrum, not duration. The model predicts that amount of L2 experience should not affect the production of similar vowels, and it was surprising to find that experience may apparently have an adverse effect on similar vowel production. The acoustic results for English /ɛ/ and /æ/ as produced by GA, and acoustic comparisons of German /ɛ/ to English /ɛ/, pointed to an explanation for this unexpected finding. While the GB subjects produced an /ɛ/ in English that did not differ significantly from their German /ɛ/, the GA subjects produced an English /ɛ/ that differed significantly from their German /ɛ/. The direction of this difference suggested that the realization rules used to output English /ɛ/ were "deflected" by the neighboring new vowel /æ/, for which they had not yet established a phonetic category.

The hypothesized influence of a new vowel on the production of a similar L2 vowel has been observed previously (Flege, 1987; Major, 1987). Major studied the production of the new English vowel /æ/ and the similar English vowel /ɛ/ by native speakers of Brazilian Portuguese. He observed that speakers who produced highly identifiable /ɛ/ tokens scored low on their intended /æ/ productions, whereas speakers whose intended /æ/ productions were highly intelligible produced /ɛ/ tokens that were frequently misidentified. Major's findings for L1 Portuguese speakers, which parallel those of the present study, are consistent with the assumption that the new vowel /æ/ may exert an influence on the similar /ɛ/ vowel as long as a category for /æ/ has not been established.

Differing from the findings of Major, and of the present study, Flege (1987) found that deflection by a new vowel did not lead to nativelike production values for the neighboring similar vowel. Instead, it caused the similar vowel to be produced farther away from the acoustic norm of the L2 than would have been the case had the nonnative speakers simply used their native vowel. Flege (1987) examined the production of English /u/, the new French vowel /y/, and the similar French vowel /u/ by several monolingual and bilingual groups. The L1 English group least experienced in French did not approximate the French acoustic norm by producing

a more posterior /u/ in French than in English. Rather, they produced a French /u/ that was even more anterior than their English /u/, suggesting that their French /u/ was influenced by the new /y/ for which they had not yet established a category. Flege's findings suggest that only after learners have established a phonetic category for a new vowel that influenced the production of a neighboring similar vowel do they start to approximate the acoustic norm for this similar vowel.

In its present version, the speech learning model does not account for possible effects of new vowels on the production of neighboring similar vowels. To accommodate the preceding interpretation of the result for English /ɛ/ (and for French /u/ in the Flege, 1987, study), the model should state that new vowels of the L2 for which phonetic categories have not yet been established may affect the realization rules that output neighboring similar vowels. Clear predictions can be made on whether the influence of a neighboring new vowel causes the similar vowel to be produced closer to the L2 acoustic norm (as for English /ɛ/ in the present study) or farther away from it. L2 speakers will approximate the acoustic norm for a similar vowel if the new and the similar vowel have common acoustic properties with respect to the similar L1 vowel (as the relatively low and relatively long English /ɛ/ and /æ/ have vis-à-vis the relatively high and relatively short German /ɛ/). On the other hand, the similar vowel will be produced farther away from the L2 acoustic norm if the new and the similar vowel in the L2 are arranged, along one or more parameters, on opposite sides of the similar L1 vowel (as the anterior French /y/ and the posterior French /u/ vis-à-vis the fairly central English /u/).

Finally, the present results indicate that English /ɛ/ is a problematic vowel for L1 Germans, even though it is not recognized as such in textbooks comparing English and German vowels. The learning problem for /ɛ/ revealed in this study is a serious one, not just because it affects intelligibility, but also because L1 Germans will apparently produce an English-like /ɛ/ only if they have not yet learned to produce an acoustic contrast between /ɛ/ and the new vowel /æ/. Once they have learned to produce this contrast, they will use a German-accented /ɛ/ in English that has acoustic properties (short in duration, high in the acoustic vowel space) that may cause its misperception as /I/. The results for GB indicate that continued L2 experience does not lead to a further approximation of the acoustic norm for the similar English vowel /ɛ/.

The prediction for the English vowel /æ/, which was shown to be acoustically quite unlike any German vowel, was that extended L2 experience will enable at least some adult L2 learners to achieve the acoustic norm for this new vowel. While the acoustic results for /æ/ were largely consistent with the prediction, the intelligibility judgments were not. The lack of a significant difference for /æ/ intelligibility between GA and GB should probably not be regarded as questioning the validity of the model, which predicted that new sounds will be produced more intelligibly by experienced than inexperienced L2 speakers. The unexpected result for /æ/ intelligibility was probably due, on the one hand, to the relatively large number of GA speakers whose intended /æ/s were correctly identified at approximately chance level because their /æ/s and /ɛ/s did not differ much acoustically.

The unexpectedly poor intelligibility of the GB speakers, on the other hand, remains unexplained in the face of their nativelike acoustic /æ/ values. We can only state that in this study, as in earlier studies of L2 speech learning (e.g., Barry, 1977), categorical judgments by native listeners failed to "honor" acoustic evidence of L2 speakers' successful efforts to sound nativelike.

It is well known to the extent of being a popular stereotype (Barry, 1981) that native Germans tend to substitute an /ɛ/-like vowel for the new English vowel /æ/ (Arnold & Hansen, 1968; Barry, 1977; Gimson, 1970; Jones, 1960; Keutsch, 1974; Kufner, 1971; Weiher, 1975; Wode, 1978, 1980, 1981). The present study showed that inexperienced native German speakers of English did not simply substitute their German /ɛ/ for English /æ/. Rather, they produced a vowel that is lower in the acoustic vowel space and longer than their native /ɛ/ for both English /ɛ/ and /æ/. This suggests that even a small amount of L2 experience with a new vowel may affect how existing phonetic categories are realized in production.

Two previous studies on the acquisition of English /æ/ in L1 (Lieberman & Kubaska, 1979) and in L2 (Barry, 1977) have indicated that learners may produce a duration contrast between /ɛ/ and /æ/ before they differentiate these vowels spectrally. In a recent perceptual study, Bohn and Flege (1990) found that inexperienced native German speakers of English labeled members from a synthetic *bet–bat* continuum primarily on the basis of vowel duration, whereas experienced Germans and particularly native English listeners based their responses more on spectral differences. These findings might lead one to expect that duration differences between spectrally neighboring vowels are acquired before spectral differences. The present study does not provide support for the existence of such a learning hierarchy in production. The speakers in GA produced significantly smaller contrasts between /ɛ/ and /æ/ than those in GB and in EN in both the spectral *and* the temporal domain. The discrepancy between these results and those reported earlier makes it clear that more research is needed to test whether or not cue hierarchies in speech learning exist.

Finally, the present results for /æ/ reaffirm previous findings (e.g., Weiher, 1975) that the targetlike production of English /æ/ constitutes a major learning problem for native German speakers. Recall that the subjects in the group labeled *inexperienced* had studied English in school for an average of over 6 years and that they had been in an English-speaking environment for more than half a year on the average. Only the experienced Germans, who differed from the inexperienced group primarily in that they had been exposed to authentic English for more than 7 years, had learned to produce an /æ/ that was much like the /æ/ of the native English speakers. This suggests that adults may need L2 exposure for a considerable period of time before they show evidence of phonetic learning for a new L2 vowel. Even then, however, regularities of the L1 may influence the production of a new vowel. Weiss's (1976) observation that "many Germans find it difficult to lengthen an open . . . vowel" (p. 12) because of the organizing principle of "long-close vs. short-open" (Weiss, 1972, p. 634) for German monophthongs may be related to the present finding that those Germans whose /æ/ productions were spectrally nativelike did

not produce sufficiently long /æ/s to match the English acoustic norm for absolute /æ/ durations.

To conclude, this study provided further evidence that the success or failure to learn L2 sounds is largely explicable in terms of sound correspondences between the L1 and the L2, and in terms of L2 experience. In particular, it was shown that differences in amount of English language experience did not affect how native German speakers produced the similar English vowels /i/ and /I/. These results, and also those for the similar English /ɛ/ vowel, were consistent with the hypothesis that extended L2 experience would not lead to the establishment of phonetic categories for similar L2 sounds because category formation is blocked by equivalence classification. Probably the most important finding was obtained for /æ/, which experienced, but not inexperienced, Germans produced close enough to the English acoustic norm to warrant the conclusion that, given sufficient L2 experience, adults can learn to produce at least some important phonetic characteristics of a new vowel category.

(Received 11 March 1991)

NOTES

1. The term *English* as used in this paper refers to American English unless stated otherwise.

2. Disner (1983) compared English /ɛ/ to German /ɛ:/ and found lower F1 values for German /ɛ:/ than for English /ɛ/ but no difference for F2 and F3. The difference for F1, however, and the large variance for F1 of German /ɛ:/ reported by Disner may have been due to North German /e:/ rather than /ɛ:/ pronunciations of some of Jørgensen's (1969) subjects, whose data Disner analyzed.

3. One subject in each German group reported that they were not currently using German, but all others used German daily at the time of the study. The mean self-estimated percentage of daily use of English was 66% for GA and 87% for GB. Because it has been suggested that listeners from southern Germany make greater use of duration in contrasting German /i–I/ than North Germans, it would have been desirable to have dialectally homogeneous German subject groups. However, due to limited German subject availability in the Birmingham (AL) region, the Germans in the present study came from various German areas. Judging from their origins, they represented the following broad dialectal backgrounds, namely, North German (GA: 3, GB: 3), Central German (GA: 5, GB: 4), and South German (GA: 2, GB: 3).

4. Tables presenting durations and untransformed frequency values of the vowels analyzed in the present study are available from the first author upon request.

5. The comparison of northern and southern German subjects revealed one difference. The two subgroups differed significantly in vowel height (B1-B0) for /ɛ:/, $F(1, 9) = 9.03$, $p < .05$, and in vowel height contrast between /ɛ/ and /ɛ:/, $F(1, 9) = 7.45$, $p < .05$. As expected, the North Germans' /ɛ:/ was higher than the South Germans' /ɛ:/.

6. The productions of EN were compared separately to those of the GA and GB so that comparisons would be based on an equal number of observations (10 subjects per group). Because both German groups differed in the same ways from EN, only the results from the comparison of the German vowels produced by GA to the English vowels produced by EN are reported.

7. The English and German data were obtained from the German subjects on separate days.

8. The B1-B0 scores of English /ɛ/ ($M = 4.28$) for GA was significantly larger than for their German /ɛ/ ($M = 3.46$), $F(1, 18) = 16.40$, $p < .001$, and /ɛ:/ ($M = 3.48$), $F(1, 18) = 13.01$, $p < .01$, and the B2-B1 score for their English /ɛ/ ($M = 6.64$) was significantly smaller than for their German /ɛ/ ($M = 7.43$), $F(1, 18) = 8.69$, $p < .01$, and /ɛ:/ ($M = 7.99$), $F(1, 18) = 23.61$, $p < .001$.

9. An alternative explanation for this result is that English language experience led the GB subjects to produce an /ɛ/ that is polarized away from the new /æ/ vowel. However, while only GB subjects produced *bet* tokens that were frequently labeled as *bit*, inspection of individual subject data failed to show a relationship between correct identification of intended *bat* tokens and misidentifications of *bet* as *bit*.

REFERENCES

Anderson, S. R. (1974). *The organization of phonology*. New York: Academic.

Anisfield, M., Anisfield, E., & Semogas, R. (1969). Cross-influences between the phonological systems of Lithuanian–English bilinguals. *Journal of Verbal Learning and Verbal Behavior, 8*, 257–261.

Anisfield, M., & Gordon, M. (1971). An effect of one German-language course on English. *Language and Speech, 14*, 289–292.

Arndt, H., & Careless, B. J. (1978). Zur auditiven Perzeption von amerikanischem Englisch durch deutsche Schüler. In H. P. Kelz (Ed.), *Phonetische Grundlagen der Ausspracheschulung* (Vol. 2, pp. 45–62). Hamburg: Buske.

Arnold, R., & Hansen, K. (1968). *Phonetik der englischen Sprache: Eine Einführung*. Leipzig: Enzyklopädie.

Barry, W. J. (1974a). Language background and the perception of foreign accent. *Journal of Phonetics, 2*, 65–89.

Barry, W. J. (1974b). *Perzeption und Produktion im sub-phonemischen Bereich. Eine kontrastive Untersuchung an intersprachlichen Minimalpaaren des Deutschen und Englischen*. Tübingen: Niemeyer.

Barry, W. J. (1977). *Ausspracheabweichungen: Eine experimentalphonetische Untersuchung an deutschen Englischlernenden* (Working Papers No. 7). Kiel: Kiel University, Phonetics Department.

Barry, W. J. (1981). Die Problematik der Aussprache in der Fremdsprachenerlernung: Die Unvermeidbarkeit eines fremden Akzents. *Working Papers, 16*, 2–28. Kiel: Kiel University, Phonetics Department.

Bethge, W. (1963). Beziehungen der Generation zur Quantität in den deutschen Mundarten. *Phonetica, 9*, 200–208.

Boesch, B. (1957). *Die Aussprache des Hochdeutschen in der Schweiz*. Zürich: Schweizer Spiegel.

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.

Delattre, P. (1964). Comparing the vocalic features of English, German, Spanish and French. *International Review of Applied Linguistics, 2*, 71–97.

Disner, S. F. (1983). *Vowel quality: The relationship between universal and language specific factors* (UCLA Working Papers in Phonetics No. 58). Los Angeles: University of California at Los Angeles, Linguistics Department.

Flege, J. E. (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, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders* (pp. 224–401). Norwood, NJ: Ablex.

Flege, J. E. (1991). The intelligibility of English vowels spoken by British and Dutch talkers. In R. Kent (Ed.), *Intelligibility in speech disorders* (pp. 157–231). Amsterdam: Benjamins.

Flege, J. E., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Flege, J. E., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

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

Gimson, A. C. (1970). *An introduction to the pronunciation of English* (2nd ed.). London: Edward Arnold.

Harris, M. S., & Umeda, N. (1974). Effect of speaking mode on temporal factors in speech: Vowel duration. *Journal of the Acoustical Society of America, 56*, 1016–1018.

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America, 33*, 1174–1178.

Iivonen, A. (1983). Zur Frage der regionalen Variation der hochdeutschen Vokale. *Neuphilologische Mitteilungen, 84*, 45–52.

Iivonen, A. (1987). Monophthonge des gehobenen Wienerdeutsch. *Folia Linguistica, 21*, 293–336.

Jones, D. (1960). *An outline of English phonetics* (9th ed.). London: Heffer.

Jongman, A., Fourakis, M., & Sereno, J. A. (1989). An acoustic study of vowels in modern Greek and German. *Language and Speech, 32*, 221–248.

Jørgensen, H. P. (1969). Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantenfrequenzen. *Phonetica, 9*, 217–245.

Keutsch, M. (1974). *Praxis der englischen Aussprache*. Tübingen: Niemeyer.

Klatt, D. H. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America, 54*, 1102–1104.

Kohler, K. (1971). On the adequacy of phonological theories for contrastive studies. In G. Nickel (Ed.), *Papers in contrastive linguistics* (pp. 83–88). Cambridge: Cambridge University Press.

Kohler, K. J. (1977). *Einführung in die Phonetik des Deutschen*. Berlin: Schmidt.

Kohler, K. (1981). Contrastive phonology and the acquisition of phonetic skills. *Phonetica, 38*, 213–226.
Kufner, H. L. (1971). *Kontrastive Phonologie, Deutsch–Englisch*. Stuttgart: Klett.
Ladefoged, P. (1978). Phonetic differences within and between languages. *UCLA Working Papers in Phonetics, 41*, 32–40.
Leskien, A., & Brugman, K. (1882). *Litauische Volkslieder und Märchen aus dem preußischen und russischen Litauen*. Straßburg: Trübner.
Lieberman, P., & Kubaska, C. (1979). Intrinsic vowel duration and formant frequencies: Data from speech acquisition. In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers* (pp. 213–215). New York: Acoustical Society of America.
Major, R. C. (1987). Phonological similarity, markedness, and rate of L2 acquisition. *Studies in Second Language Acquisition, 9*, 63–82.
Meyer, E. A. (1904). Zur Vokaldauer im Deutschen. In *Nordiska Studier tillegnade Adolf Noreen* (pp. 347–356). Uppsala: Appelberg.
Miller, J. D. (1989). Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America, 85*, 2114–2134.
Moulton, W. G. (1962). *The sounds of English and German*. Chicago: University of Chicago Press.
Moulton, W. G. (1987). Changing standards for Standard German pronunciation and some student reactions. *Beiträge zur Phonetik und Linguistik, 52*, 261–289.
Narahara, Y., Okamoto, T., & Shimoda, H. (1977). *Beiträge zur Phonetik des Deutschen*. Hamburg: Buske.
Oakeshott-Taylor, J. (1976). Perception and production of English sounds by German speaking subjects. *Hamburger Phonetische Beiträge, 17*, 163–199.
Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184.
Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.
Rietveld, A. C. M. (1975). Untersuchung zur Vokaldauer im Deutschen. *Phonetica, 31*, 248–258.
Scherer, G., & Wollmann, A. (1977). *Englische Phonetik und Phonologie* (2nd ed.). Berlin: Schmidt.
Schouten, M. E. H. (1977). Imitation of synthetic vowels by bilinguals. *Journal of Phonetics, 5*, 273–283.
Syrdal, A. K. (1985). Aspects of a model of the auditory representation of American English vowels. *Speech Communication, 4*, 121–135.
Syrdal, A. 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.
Traunmüller, H. (1982). Der Vokalismus des Ostmittelbairischen. *Zeitschrift für Dialektologie und Linguistik, 49*, 289–333.
Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 58*, 434–445.
Weiher, E. (1975). *Lautwahrnehmung und Lautproduktion im Englischunterricht für Deutsche*. Unpublished doctoral dissertation, Kiel University, Kiel.
Weiss, R. (1972). Perceptual parameters of vowel duration and quality in German. In A. Rigault & P. Charbonneau (Eds.), *Proceedings of the Seventh International Congress of Phonetic Sciences* (pp. 633–636). The Hague: Mouton.
Weiss, R. (1976). *The perception of vowel length and quality in German: An experimental-phonetic investigation*. Hamburg: Buske.
Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception & Psychophysics, 26*, 95–104.
Wode, H. (1978). The beginnings of non-school room L2 phonological acquisition. *International Review of Applied Linguistics, 16*, 109–125.
Wode, H. (1980). Phonology in L2 acquisition. In S. W. Felix (Ed.), *Second language development: Trends and issues* (pp. 123–136). Tübingen: Narr.
Wode, H. (1981). *Learning a second language: Vol. 1. An integrated view of language acquisition*. Tübingen: Narr.
Wurzel, W. U. (1984). Phonologie: Segmentale Struktur. In K. E. Heidolph, W. Flämig, & W. Motsch (Eds.), *Grundzüge einer deutschen Grammatik* (pp. 898–990). Berlin: Akademie-Verlag.