

Detection of Foreign Accentedness

J. E. Flege
Birmingham, U.S.A.

1. Introduction

Humans derive several important kinds of information from the speech signal in addition to referential meaning. One of these is 'accentedness,' a realization by the listener that his interlocutor differs in social, geographical, or linguistic background. This study examines the ability of American listeners to detect accent in the speech of French-English bilingual talkers. Previous studies have revealed two important aspects of accentedness. First, it lessens intelligibility, at least in non-optimal (e.g. noisy) conditions (Lane, 1967; Lehtonen and Leppanen, 1980). Native speakers can identify more of the words produced by fellow native speakers than by non-native speakers. Similarly, the non-native speaker is able to identify more words produced by native speakers of the target language than by other non-natives, including those who share the same mother tongue (Johansson, 1978). As children mature they become better able to detect departures from the phonetic norms which uniquely characterize their native dialect (Elliot, 1979; Scovel, 1981). The ability of L2 learners to detect accent and to authentically pronounce the target language gradually improves, although L2 learners may never match native speakers of L2 in either respect (Flege, 1980; Flege and Port, 1981; Scovel, 1981). All of these developments suggest a slow evolution in the internal phonetic representation of speech sounds and the ability to detect departures from them.

2. General Method

English phrases were produced by 8 monolingual native speakers of American English and 8 native speakers of French, all women. The French speakers were women 28-48 years of age who had lived in Chicago an average of 13 years at the time of the study and who all spoke English with an obvious French accent in the author's estimation. No attempt was made to control the native dialect of the French speakers. Four were from Paris, two from Belgium, one from Annecy, and one from St. Etienne. The English phrases examined in this study were either read from a list or produced in utterance initial position during a spontaneous story. It was hypothesized that the second, more demanding task would result in less authentic English pronun-

ciation by the French speakers since it presumably required greater attention to the content of the speech being produced as opposed simply to its phonetic form.

A waveform editing program was used to isolate increasingly shorter portions of the digitized speech signals and to cross splice portions of words. Stimuli in the five related experiments to be reported here were normalized for RMS intensity before being converted to analog form with 12-bit resolution, filtered, and presented to subjects (listeners) binaurally at a comfortable listening level. The subjects, all native speakers of American English, were seated in a sound booth before a response box with two buttons. Stimulus randomization, presentation, and data collection were all run under the real-time control of a small laboratory computer (PDP 11/34).

2.1. Experiment I

Two replicate productions by each speaker of the phrases 'Two little birds' and 'Two little dogs' from both of the two speaking conditions (Isolated Phrase, Spontaneous Story) were presented separately for forced choice identification as 'native' (American) or 'non-native' (French). The subjects in this experiment were 10 Americans with a mean age of 32 years who had backgrounds in phonetics, linguistics, or French. Of these subjects, 3 spoke French and 4 indicated 'frequent' exposure to French-accented English.

The subjects were easily able to identify the phrases produced by the non-native speakers as 'non-native.' There was no difference between speaking conditions. In the Isolated Phrase condition the French women were correctly chosen as 'non-native' 91% (1087) of the time; in the Story condition 88% of the time (1056 judgments). There were less than 1% incorrect identifications of American women as 'non-native.'

The phrases produced by most of the individual French Speakers were correctly identified more than 90% of the time. One French speaker's production of 'Two little dogs' was accepted as 'native,' but it should be noted that her language background was correctly identified in 'Two little birds,' probably due to the deletion of /r/ in 'birds.' This suggests that a specific sound may continue to pose difficulty for a language learner, even in the face of a generally close approximation to the phonetic norms of a foreign language. It also suggests that the 'distortion' (or omission) of a single phonetic 'segment' may cue foreign accent.

2.2. Experiment II

Next, the first syllables of the phrase length stimuli used in the first experiment (/tu/) were presented to subjects for paired comparisons. The subjects' task was to determine which member of the pair had been produced by the 'non-native' (French) speaker. The 10 subjects (mean age, 28 years) were sophisticated in that they had training in phonetics or spoke French.

Overall, the subjects correctly chose the /tu/ produced by the French speakers as 'non-native' 95% (5447) of the time. Syllables produced by the Americans were incorrectly chosen only 5% (313) of the time. There was again no difference between the Isolated Phrase and Story conditions, so the remaining experiments examined only speech taken from the Isolated Phrase condition.

2.3. Experiment III

Several measures were taken to generalize the findings of experiments I and II. The 10 subjects (mean age 22 years) in this experiment were students having little or no experience with French or French-accented English. In addition to /tu/ the /ti/ syllable from 'TV reception' and 'TV antennas' was also presented for identification as 'native' or 'non-native.'

Overall, syllables produced by the French speakers were correctly identified 76.7% (2945) of the time. The Americans were incorrectly identified as 'non-native' only 22.3% (857) of the time. Both the frequency of correct identifications and rejections differed significantly from chance ($p < .01$).

Systematic debriefing after the experiment revealed that subjects could not identify the native language of the French speakers. Thus their ability to detect accent probably resulted from sensitivity to departures from English phonetic norms rather than a tacit knowledge of the phonetic characteristics of French-accented English.

2.4. Experiment IV

The phonetic differences that might cue accent were now restricted to the domain of a single phonetic 'segment' (i.e., /t/, /u/, or /i/). One set of hybrid syllables was created by electronically splicing the many /t/'s produced by native and non-native speakers onto a single, good exemplar of an English vowel (/i/ or /u/). Another set was created by splicing the multiple /i/ and /u/ vowels produced by native and non-native speakers onto a single good exemplar of English /t/ (edited from /ti/ and /tu/ syllables, respectively). The American-American and French-American hybrid stimuli were presented in pairs. The subjects' task was once again to identify which of the two stimuli per trial sounded 'non-native.'

Differences sufficient to cue accent resided in just the /t/, /i/, and /u/ segments. The correct recognition rates were: 63% (/t/ + constant /i/); 71% (/t/ + constant /u/); 66% (constant /t/ + /i/); and 69% (constant /t/ + /u/). All were significantly above chance levels ($p < .01$).

2.5. Experiment V

The paired comparison method was again used in the final experiment to examine whether differences in just one portion of a phonetic 'segment'

might cue accent. The first 30 ms (± 1 ms) of /tu/ was presented to 9 students in Speech-Language Pathology, all of whom had some training in phonetics. Although these '/t/-burst' stimuli were not recognizable as speech sounds, subjects were instructed to use their own pronunciation of /t/ as a standard for determining which of the two stimuli per trial had been produced by a non-native speaker. No other training feedback, or familiarization was given.

In 68.7% (3164) of the cases the /t/-bursts produced by the French speakers were correctly identified as 'non-native' ($p < .01$). Eight of the 9 subjects performed at above-chance rates. Of the 8 French talkers, 6 were identified above chance ($p < .01$).

3. Discussion

This study shows that listeners are able to detect accent on the basis of small differences in the quality of /u/ or /i/, or in the place of articulation for /t/. One might simply assume this finding to be inevitable, since children learning English and French as native languages learn to produce /i/, /u/, and /t/ according to the phonetic norms of those languages. However, the acquisition process probably never involves the direct comparison of the kind of small phonetic differences that distinguished the native French and English speakers in this study. Our subjects might also have learned to 'filter out' the phonetic differences that distinguish cognate sounds in French and English since they are not relevant to phoneme identity in either language. These results demonstrate that adults do possess the ability to detect within-category (subphonemic) differences between language varieties, and to use this information in detecting foreign accent.

The present results suggest that listeners develop phonetic category prototypes against which to judge specific speech sounds. Although we have no direct evidence concerning the articulatory differences distinguishing the /t/ produced by the native French and English talkers, it is mostly likely a difference in the width and place of tongue contact. The French /t/ was sufficiently far from the range of auditory properties acceptable as /t/ that it was rejected as an English sound. An important question for future research is how 'accentual' information and the information cuing category identity are processed during speech perception. A recent study indicated that subjects' responses to stimuli changed as a function of whether the stimuli were presented as 'Spanish' or 'English' (Elman, Diehl and Buckwald, 1977). A 'prototype' model of accent detection fails to predict such a finding since it presumes that information relevant to phonetic category identity must be processed prior to assessment of 'accent' (i.e., departures from phonetic norms for a particular sound category). One possibility raised by this finding is that accentual information is processed integrally along with information leading to the identification of a phonetic category.

References

- Elliot, L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using material with controlled word predictability. *J. Acoust. Soc. America*, **66**, 651-653.
- Elman, J., Diehl, R. and Buchwald, S. (1977). Perceptual switching in bilinguals. *J. Acoust. Soc. America*, **62**, 971-974.
- Flege, J. (1980). Phonetic approximation in second language acquisition. *Lang. Learning*, **30**, 117-134.
- Flege, J. and Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Lang. and Speech*, **24**, 125-146.
- Johansson, S. (1978). *Studies in Error Gravity*. Gothenberg Studies in English, **44**. Department of English, Univ. of Gothenberg.
- Lane, H. (1963). Foreign accent and speech distortion. *J. Acoust. Soc. America*, **35**, 451-453.
- Lehtonen, J. and Leppanen, H. (1980). Observations on the effect of background noise on the understanding of English by Finnish students. *Papers on Speech Res.*, **2**, 108-116. Inst. Finnish Lang. and Communication, Univ. of Jyvaskyla.
- Scovel, T. (1981). The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition. In: J.-G. Savard and L. Laforge (eds.). *Proceed. 5th Int. Cong. Int. Assoc. of Applied Ling.* Quebec: Laval Univ. Press.