PERCEPTION OF SYNCOPE IN NATIVE AND NON-NATIVE AMERICAN ENGLISH

Joann Fokes and Z. S. Bond

Ohio University, Athens, Ohio, U.S.A.

ABSTRACT

Native and non-native English speaking subjects made forced choice identifications of word triads embedded in phrases as spoken by three different English speakers. The triads consisted of 1) words with initial unstressed [as] syllables, 2) words created by vowel syncope resulting in s-clusters, and 3) words containing s-clusters. A three way analysis of variance revealed a significant interaction between the two subject groups, word triads, and the speakers. Native subjects were better able than the non-natives in identifying tokens even though there were no differential peaks in production. There was some bias in terms of speaker and particular word stimuli.

1. INTRODUCTION

Both native and non-native speakers alter the pronunciation of English in casual speech, but perhaps in different ways. For example, native Americans frequently employ syncope or vowel loss in the pronunciation of unstressed syllables. This phenomenon is well documented [3] in the case of internal unstressed syllables and appears to be correlated with word stress patterns. Such reductions seem to be more common in English than other languages because of its polysyllabic rhythm. Typically, syllables containing strong beats fall at irregular intervals and are surrounded or flanked by syllables with weak beats. Reductions also occur in initial unstressed syllables as in the casual pronunciation of close for suppose. In fact, vowel syncope may spill over into the normal styles as in the network news commentary reporting recent Supreme Court decisions: "S'press Court decisions".

In the preceding example, vowel syncope results in a word with two juxtapositioned consonants resembling a dictionary word which does indeed contain a cluster. For example, vowel syncope occurs in the production of sport which then becomes a possible homonym with sport. Just how listeners identify words containing vowel loss which become homonyms with real words is the question of interest in this investigation. It can be hypothesized that correct word identification is based on the semantic content of the message. On the other hand, there could be conclusions in the perception of the target word unless the phonetic characteristics of the utterance provide for cues in its correct perception. Thus, if the content is ambiguous, there could be phonetic information to aid in the perception of the intended word.

Before the perception of words containing vowel syncope can be adequately studied, the actual production of such items requires description. The phonetic detail of clusters resulting from vowel syncope was previously investigated by Fokes and Bond [4,5]. They tape recorded ten American English speaking subjects and four non-native English speakers who read a series of six phrases or sentence sets. Each set contained a triad of test words embedded in the same phrase: 1) a word beginning with an unstressed syllable in the form of [as] followed by [sp] or [sk]; 2) a word containing an initial cluster consisting of [sp] or [sk], and 3) a word containing a real cluster resulting from vowel syncope. The subjects reported no difficulty in differentiating words with unstressed syllables, clusters, and initial unstressed syllables even though there were no differential peaks in production. There was some bias in terms of subject and particular word stimuli.

2. METHOD

2.1. Materials

The stimuli for the present study were the productions from the previous investigation and consisted of tapes recorded readings of short phrase or sentence triads containing test words 1) with an initial unstressed [as] syllable, 2) a real [sp] or [sk] cluster, and 3) an artificial [sp] or [sk] cluster.
ter. Each member of a triad was inserted into the following phrase sets:

On (sucumbing, succumbing, succumbed) parties.
He (secured, skewered, cured) the meat.
The (supplies, splice, splice) of tape.
My (support, sport, sport) of baseball.

Four tokens of each item spoken by three native Americans and one proficient non-native speaker who had been speaking English since childhood were recorded in random order to make a listening tape of 192 items. The speakers were selected on the basis of clarity of the tape and the absence of any trace of an unstrained vowel in words containing either the artificial or real clusters. The reduced vowel was present in the two-syllable words.

2.2. Subjects
The two groups of subjects were college students: 15 native American English listeners and 10 non-native listeners. The non-native group’s experience with English was limited to academic training in English in their homeland and from two to five years contact at Ohio University.

2.3. Procedure
The subjects made forced-choice identifications (ex: splice/supply) of each of the tape-recorded tokens. Subjects listened via head-phones in a quiet listening laboratory.

3. RESULTS
The percent identifications of the triads by both groups of listeners are given in Table 1. The American listeners identified real clusters and two-syllable words nearly 100% of the time. They heard the artificial clusters as two-syllable words at variable rates ranging from 56.4% for one of the native productions to only 7% for the non-native proficient speaker.

Non-native listener identifications of real clusters ranged from 79% to 98% and from 65% to 96% for two-syllable words. They identified artificial clusters as two-syllable words from 51% for the non-native speaker to 47% for 48% of the native speakers. Interestingly, the non-native listeners perceived the proficient non-native speaker’s artificial clusters as the target word more often than the native subjects.

Identification scores were also lexically dependent; g’cumb was rarely heard as g’cumb (8%), while sport and s’port were identified as two-syllable words 64% of the time. In fact, with the word agriculture removed from the analysis, the test word of the artificial cluster rose to 56% for speaker four’s productions and to 65% forspeaker two. Identification scores also rose to a level of 38% for the non-native speaker productions as well.

Identification scores were submitted to a 2 by 3 by 4 repeated measures analysis of variance, consisting of one between factor (two listener groups), and two within factors (4 English words and two triads). The Greenhouse-Geisser adjusted degrees of freedom were used to test the interaction and main effects. There were the following significant interactions: speaker by listener group (F = 4.74; df = 2.16; p < 0.01); speaker by word triad (F = 36.11; df = 3.26; p < 0.0001); and speaker by word triad (F = 5.81; df = 3.26; p < 0.0009). There was no listener group by word triad interaction. In determining these interactions, Speaker one was clearly different in that her artificial clusters could not be identified as intended by native subjects but were identified at somewhat higher rates by non-native subjects. Also significant were the main effects of listener group (F = 23.35; df = 1, 23; p < 0.0001), (speaker (F = 3.27; df = 2, 27; p < 0.01) and word triads (F = 64.22; df = 1.44; p < 0.0001).

4. CONCLUSIONS
The American native subjects were better able to identify artificial clusters as the target word containing a vowel not intended by the speakers than the non-natives. This ability cannot be credited to semantic cues on the intention of the speaker because the words were embedded in the same phrase. Subjects, however, were highly influenced by specific words and the linguistic background of the speaker.

Because there was no single invariant acoustic pattern separating real from artificial clusters, we speculate that both groups of listeners were using multiple cues as a basis for perceptual judgments. That is, one speaker may have used a set of cues which, in turn, may have signaled the intended target word. Differences in listener judgments may have the facility of adapting to the peculiarities of individual speakers and the linguistic environments listeners are using in word-initial consonant clusters, Journal of Speech and Hearing Research, 18, 666-706.

| Table 1. Mean and 95% confidence intervals for native and non-native English subjects in identifying the stimulus triads. |

<table>
<thead>
<tr>
<th>Triad Cluster</th>
<th>Native Mean</th>
<th>95% CI</th>
<th>Non-native Mean</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lace</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>92.2</td>
<td>91.9-92.5</td>
<td>90.4</td>
<td>89.8-91.0</td>
</tr>
<tr>
<td>2</td>
<td>90.7</td>
<td>90.3-91.1</td>
<td>89.1</td>
<td>88.6-90.1</td>
</tr>
<tr>
<td>3</td>
<td>98.7</td>
<td>98.3-99.1</td>
<td>97.2</td>
<td>96.7-97.6</td>
</tr>
<tr>
<td>4</td>
<td>98.2</td>
<td>97.9-98.5</td>
<td>97.1</td>
<td>96.6-97.5</td>
</tr>
</tbody>
</table>

5. REFERENCES