Acoustic Structure of English Diphthongs and Semi-Vowels vis-à-vis Their Phonemic Symbolization

By C. M. Wise, Baton Rouge, Louisiana

What are the terminal vowel sounds of English diphthongs? Are the transient sounds between the initial and terminal vowels significant? Chao\(^1\) finds that they are not. Accordingly the terminal vowel of a diphthong may be examined in isolation to discover what it is.

Consider the diphthong in *bite* (Fig. No. 1). The second spectrogram has been gated by an electronic switch to eliminate 30 ms of the transition portion of the diphthong just before the steady state of the terminal vowel. The elimination is scarcely perceptible, and

---

*Fig. 1. a Spectrogram of bite without modification. b Spectrogram of bite ‘gated’ with a gap of 30 ms.*

* Credit due to André-Pierre Benguerel for laboratory work.
there is no phonemic change. It may be concluded that the portion removed is not essential to the hearer's interpretation of the word.

By means of the dual-loop recorder with associated spectrograph and gate circuit, the whole transition can be eliminated and the initial and final parts moved together (Fig. No. 2). The word is now shorter, but its meaning is the same, and the diphthong as such is not affected.

The foregoing process has been applied also to the diphthongs in *bait, boat, bout* and *boat,* with analogous results.

Speech is a phenomenon which human beings can hear as "a sequence of auditory fractions" or discrete elements, so "structured" as to "permit the exclusion of some features of the speech continuum from linguistic consideration." By such exclusion of the transition within a diphthong it is again found possible to isolate end-vowels to find out what they are.

We now examine a spectrogram of the word *bile* with the end-vowel of the diphthong isolated (Fig. No. 3). With the gate circuit set to pass only this vowel, we record it. In the same way, we record the end-vowel of the diphthong in *boat.*

These brief utterances are hard to identify. Listeners often adjudge the diphthong of *bile* as ending in either /I/ or /i/, and of *boat* in /u/ or /w/. We now test these judgments by the measurement of the first and second formant frequencies of the two vowels and by

<table>
<thead>
<tr>
<th>Informant CMW</th>
<th>Vowel</th>
<th>F1 (kHz)</th>
<th>F2 (kHz)</th>
<th>Vowel</th>
<th>F1 (kHz)</th>
<th>F2 (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bite</td>
<td>/aV</td>
<td>435</td>
<td>2550</td>
<td>/I</td>
<td>275</td>
<td>2550</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>420</td>
<td>2350</td>
<td>/i</td>
<td>200</td>
<td>2700</td>
</tr>
<tr>
<td></td>
<td>/aV1</td>
<td>280</td>
<td>2145</td>
<td>/I</td>
<td>275</td>
<td>2350</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>450</td>
<td>2390</td>
<td>/i</td>
<td>200</td>
<td>2700</td>
</tr>
<tr>
<td>bait</td>
<td>/aV</td>
<td>410</td>
<td>2230</td>
<td>/I</td>
<td>275</td>
<td>2350</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>450</td>
<td>2190</td>
<td>/i</td>
<td>200</td>
<td>2700</td>
</tr>
<tr>
<td>boat</td>
<td>/aV</td>
<td>445</td>
<td>800</td>
<td>/I</td>
<td>375</td>
<td>975</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>450</td>
<td>1000</td>
<td>/i</td>
<td>490</td>
<td>825</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>300</td>
<td>600</td>
<td>/w</td>
<td>375</td>
<td>975</td>
</tr>
<tr>
<td></td>
<td>/aV1</td>
<td>380</td>
<td>870</td>
<td>/e</td>
<td>450</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>450</td>
<td>825</td>
<td>/u</td>
<td>450</td>
<td>825</td>
</tr>
<tr>
<td></td>
<td>/aV</td>
<td>350</td>
<td>490</td>
<td>/| 380</td>
<td>885</td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 2. Spectrogram of *bite.* Glide of 115 ms indicated.

Fig. 3. Spectrogram of /I/ of *boat.*
comparing them with the formant frequencies of sustained utterances of /i-/ and /u-/ phonated by informant CMW. Here follows table I, showing the comparisons.

These comparisons show that informant CMW's end-vowels in the diphthongs of bite, bait, boit prove to be /i/ and /j/. The end-vowels of the diphthongs in bout, boat range from /u/ through /o/ to /o/. These findings invite measurements of the formant frequencies of many speakers.

It may be noted that the formant measurements of these end-vowels show no close relation to those of /j/ and /w/ respectively. The F1 of /j/ is very much lower than any value of /F1/, and F2 much higher; and both F1 and F2 of /w/ are significantly lower than those of /F2/.

Thus our evidence points toward the conclusion that the end-vowel of the diphthongs of bite, bait, boit lies in the range of /i—i/ but far from /j/; and that the end-vowel in the diphthongs of bout, boat lies in the range of /u—u—o/, but significantly far from /w/. These findings are supported by Lehiste.

Table II, omitted here, compares the formant frequencies of table I with those of the same vowels, as reported by earlier experimenters (Ilsa Lehiste, G. E. Peterson, H. K. Dunn). It confirms table I.

The use of /j/ for the end-vowel of the diphthongs in bite, bait, and /w/ for that of bout, boat appears to be based on the premise that certain recorded speech sounds give similar auditory impressions if played in reverse. Steady-state vowels, with onset and terminal transition removed, do sound about the same backward, but less so with the onset and terminal transition restored. Sounds in a sequence, such as /atʃa/ and /adʒa/, sound unlike themselves when reversed.

It is also true that yea and woe sound somewhat alike when reversed; but the two versions of each syllable are only mirrored images of each other, like the right and left hands.

The phonemes /j/ and /w/ require brief steady-state vowels as beginnings, followed by rapid transitions after them, not before them, as in the terminals of yea and woe; hence /j/ and /w/ are ruled out as terminals.

However, playing utterances in reverse sometimes reveals clues for testing the validity of suspected phonemic features. The recorded word law played in reverse sounds considerably like hall because a residue of unvoiced breath escapes after the /s/ is finished. In re-verse, this rush of air attracts attention and sounds like /h/. But it has no linguistic significance.

The foregoing suggests that a similar escape of breath has been responsible for the interpretation of the terminal sound of here as /h/. The effect of an /h/ may appear when here, pronounced /hɛr/, is played backward: but then the escaping breath precedes the series of sounds, so that /hɛr/ results. A variation of this phenomenon can be observed when cord, pronounced /kɔ:d/, sometimes transcribed /kɔd/, is played in reverse and emerges as /ɛr/ko/. This time the effect of /h/ is completely separated from the vowel by the /d/.

Those who hit upon the transcription /hɛr/ for here must have mistaken the escape of air for the final vowel of the word, whereas it was no part of the word. Since the vowel remains /ɛ/, both in the forward and the reverse playing of the word, and since, when played forward, it is a clearly defined schwa, there seems no reason for symbolizing it otherwise.

Summary

The transition between the steady-state vowels of a diphthong does not appear essential to its interpretation. The end-vowel in bite, bait, boit is phonetically in the range of /i/ or /i/. It is not /j/.

The end-vowel in bout, boat is usually in the range of /u/ or /o/ (occasionally /o/). It is not /w/.

The final phoneme of here /hɛr/, there /tɛr/, when terminal, as well as the second vowel in cord /kɔd/, is /ɛ/, not /h/. What has sometimes passed for /h/ appears to be the irrelevant expiration of breath after all the phonemes of a terminal word have been uttered.

Conclusion

It would seem that these phonetic facts should not be disregarded in the symbolization of the phonemes affected by them.

References


Author's address: Professor C. M. Wise, 4218 Tulane Drive, Baton Rouge, Louisiana (USA).