

In: Bell-Berti, F., & Raphael, L. J. (1995).
Producing Speech: Contemporary Issues. For
 Katherine Safford Harris. AIP Press: New York.

9 English /w, j/: Frictionless Approximants or Vowels Out of Place?

Leigh Lisker
Haskins Laboratories
University of Pennsylvania

The existence of strictly alphabetic representations of spoken language, of the kinds practiced by linguists, has sometimes been taken to mean that any speech signal may be described as a sequence of discrete sounds, and that a difference in the spelling of two utterances implies a dissimilarity, audible and "phonetic," that is potentially of semantic significance. Sounds spelled alike are heard to be linguistically, and perhaps also phonetically, the same. With the advent of methods of rapid spectrographic analysis some researchers have come to believe that speech cannot be analyzed into acoustically specifiable segments isomorphic with their alphabetic spellings. The linguist's spelling of a speech signal, therefore, is not equivalent to an acoustic description of its linguistically relevant properties.¹ The basis for this conclusion, not universally accepted, is

¹Nor is it very possibly isomorphic with any physiological account, despite hopes expressed by some speech researchers. We may remember that at least one linguist, not primarily a phonetician, long ago remarked that "Since the representation of an utterance or its parts is based on a comparison of utterances, it is really a representation of distinctions. It is this representation of differences which gives us discrete combinatorial elements (each representing a minimal difference). A noncomparative study of speech

that sounds spelled alike, and auditorily judged to be the same phonetically,² are always found to differ acoustically even when produced by the same vocal tract.³ Some of these differences are no doubt auditorily irrelevant, but others can be shown to be systematically related to differences in context and rate of articulation (whatever their perceptual significance might be). Those who are convinced that a purely acoustic account of speech perception is unachievable have raised the question as to whether there might not be a closer fit between the discrete model of speech as sounds in sequence and the movements in the vocal tract by which speech is generated. These movements, the "tract variables" they affect, and perhaps the motor commands that activate them, have been described as discrete and sometimes temporally overlapping (Browman & Goldstein, 1992; Harris, 1984; Liberman & Mattingly, 1985) and it is yet too early to decide whether they can serve as the elements of a description that either 1) more directly matches alphabetic perception than does one based on selected acoustic properties, or 2) replaces the alphabetic model with one that better accounts for the motor as well as some perceptual aspects of speech communication. It must be noted, however, that a model which features overlapping elements can hardly be said to serve as an explanation of the central fact of speech perception that is the basis of an alphabetic representation.⁴ Of course, the processes of speech production, whatever their relation to perception and alphabetic writing might be, are inherently interesting to those of an inquiring turn of mind. If one's purpose is to relate production to perception, it is not clear what is the "best" way to describe production,— whether in terms of electromyographically determined "motor commands," related movements of the mobile parts of the respiratory tract, or the size and shape of the cavities of that tract. With respect to articulatory activity, moreover, one may feel compelled to decide whether a dynamic description of movement trajectories or one couched more abstractly in terms of static target shapes is more revealing.⁵ Indeed, it is not clear that we have here a properly posed question, since a physiological description in terms of variously phased vocal tract movements is not strictly incompatible with one that posits vocal tract shapes as stationary

behavior would probably deal with complex continuous changes, rather than with discrete elements" (Z. S. Harris, 1951, p. 367).

²They would therefore be written alike no matter how "narrow" the representation. In broader spellings, where phonetically differentiable sounds are sometimes written alike, it is hardly surprising to find acoustically diverse sounds represented by the same letter.

³The acoustic identity of two intervals of speech would strongly suggest that they were recordings of a single phonetic event.

⁴Though, as Harris years ago pointed out in a statement that still holds true, "we appear to have failed to find a simple, absolutely invariant correlate of the phoneme, at the peripheral levels thus far investigated," (Harris, 1974, p. 2298).

⁵Browman and Goldstein (in press), for example, report that the management of lip closure for [p, b, m] varies with vocalic contexts, but that the lip aperture achieved is "relatively" context-independent.

targets toward and away from which those movements are directed. Moreover, the phonetic literature suggests the possibility, inelegant though it might be, that the same mode of description may not be equally expedient for all the phonetic elements (whether thought of as sounds or articulatory events) attributed to speech. Thus some elements may conceivably be characterized as movements without targets (e.g., the discussion of English schwa by Browman & Goldstein, 1992), while others involve targets without invariant characteristic movements.⁶

The inventory of sounds reported for English includes two vowels often now represented phonologically as /u/ and /i/. /u/ is conventionally said to be a high back rounded vowel, while /i/ is high front unrounded. Such vocal tract shapes are by definition characteristic of the IPA cardinal vowels [u] and [i], respectively, and we may suppose that for some varieties of English these IPA norms are not very far from the targets of tongue and lip movements used to produce the vowels in, for example, the words *ooze* and *ease*.⁷ It might be supposed that all the other sounds of English are somehow different from them in their articulatory properties. However, included in the English sound inventory are two other items, /w/ and /j/ (not presumably very different from IPA [wɹ]),⁸ that are often described (though not invariably, cf. Lehiste, 1964) as sounds involving vocal tract shapes identical, or very nearly identical, with those of /ui/ (= [ui]_{Eng}): /w/ is high back rounded like /u/, while /j/ is like /i/ in being high front unrounded. Indeed, in some treatments of English phonology, it has been questioned whether /w/ and /j/ should be recognized as phonemes distinct from /u/ and /i/ (Swadesh, 1947). If, despite the fact (?) that /w/ = /u/ and /j/ = /i/ with respect to the target vocal tract shapes ascribed to them, /wɹ/ are nevertheless perceived and treated as sounds differing from /ui/, then it would seem that while target identity is sometimes proposed to explain the perceptual identity of acoustically different sounds, target identity does not always guarantee perceptual invariance.

While the pairs /uɹ/ and /iɹ/ are sometimes said to be characterized by only one target shape each, a common way to differentiate /wɹ/ from /ui/ is to call them "glides," suggesting that they are "dynamic" in a way or to a degree that all other sounds are not, despite the often enough expressed observation that speech is in general produced by a vocal tract of continuously changing shape. The boundary between glides and vowels is somewhat fuzzy, since /wɹ/ may involve, at least in initial position, brief steadystate intervals (Lehiste, 1964;

⁶Thus monophthongal vowels and fricatives might be characterized by target vocal tract postures, but not with invariant movements, no matter how much information the listener may extract from the context-specific transitional intervals of connected speech.

⁷In this paper "/ui/" will refer to "[ui]_{Eng}," i.e., those pronunciations of the English phonemes /ui/ that are phonetically closest to IPA [ui]. Other realizations of /u/ and /i/, e.g., [iʷ] and [ij] (cited by Trager & Smith, 1951), will not be considered.

⁸According to Ladefoged, though, "English /w/ is in between the French sounds /w/ and /ɥ/" (Ladefoged, 1982, p. 209).

O'Connor, Gerstman, Liberman, Delattre, & Cooper, 1957), while in the vowels of fluent speech steadystate intervals are not invariably present (Lindblom & Studdert-Kennedy, 1967). Since in fluent speech transitions as surely mark vowels as they do glides, and their contribution to perception is not negligible (Nearey, 1989; Strange, 1989) then the difference seems to be that only the vowels can be produced in isolation, while glides occur only in juxtaposition to vowels.⁹ The production of /wɹj/, then, necessarily involves articulatory movements with their acoustic consequences, that is, transitions, while /ui/ can be produced with the vocal tract held fixed over an appreciable time interval, even if in speech transitions too will necessarily be present. Catford goes so far as to make the point that /wɹj/ "cannot be prolonged" (Catford, 1977, p. 128), though this seems to be only doubtfully true.¹⁰

Another basis for the distinction drawn between glides and homorganic vowels refers to a difference in degree of prominence within syllables: "...the entire difference between the English phonemes /wɹj/ and /ui/ is definable only in these terms: where vocoids of the high back rounded or high front unrounded types are the most prominent elements in syllables, they are instances of the phonemes /ui/, but where they occur as marginal elements in a syllable with something else as the most prominent part, they are instances of the phonemes /wɹj/" (Hockett, 1958, p. 82).¹¹ Such a difference would of course be prosodic in nature rather than a matter of vocal tract shaping. It should be noted that, oddly enough, this description of the phonetic basis of the contrast between /wɹj/ and /ui/ is virtually the same as the one that Swadesh (1947) used to argue that the glides do *not* constitute a phonologically distinct class of phonetic events.

Other differences between /wɹj/ and /ui/ have been adduced. Thus not everyone is agreed that these glides and vowels share precisely the same targets, though /wɹj/ are both labial-velar and /ui/ are both nonlabial-palatal. Acoustically the two sets have been said to differ in their formant structures, particularly the second formant (F₂), while in production the glides are characterized by significantly greater constrictions (Chomsky & Halle, 1968; Lehiste, 1964), even to the point of generating audible frication, at least for /j/ (Heffner, 1950). Furthermore, the targets postulated for the glides have sometimes been said to be context-determined, so that /w/ always involves movement from a position that is more high-back-rounded than that of a

⁹"Steadystate /w/ and /j/ are quite out of the question since the absence of transitions can only result in /u/ and /i/" (O'Connor et al., 1957, p. 31)

¹⁰Languages with long or geminated [w,j] are not unknown, e.g., in Cairene Arabic (I.P.A., 1949, p. 34).

¹¹For Hockett, in line with general American practice, English /ui/ are closer to IPA [uɪ] than to cardinal [ui], while the vowels more like [ui] are represented as the sequences /uw i y/. Assertions that /wy/ have the same target positions as /ui/ may therefore sometimes be understood to equate the English glides with the "lax" vowels [uɪ].

following vowel, while /j/ is a movement from a more high-front-unrounded position (Trager & Bloch, 1941, p. 234).

If the basis for calling /wj/ "glides" is that they cannot be produced and recognized as isolated utterances, then of course stop consonants must also be recognized as glides. In fact, stops are even more deservedly called glides (though they are not) since not only their intelligibility, but their status as speech sounds depends on the presence of transitions. (Given no transitions, any isolated signal emitted during a stop closure will be only doubtfully identified as speech.) On the other hand, any /wj/ steady-state intervals, if extracted and presented in isolation, are likely to be identified as vowels, most likely /ui/ or possibly /uɪ/. One difference between glides and stops is in the nature of the F₂ targets specified: the target frequencies may be achieved in the case of the glides, while for the stops the so-called "loci" are frequencies that the F₂ transitions presumably aim toward but regularly fall short of (Delattre, Liberman, & Cooper, 1955). Measurements of natural speech as well as experiments in synthesis seem to bear out this suggested difference. Thus F₂ transitions with the same terminal frequencies adjacent to different vowels do not as a rule evoke the same stop percept (Sussman, McCaffrey, & Matthews, 1991), but intervocalic /wj/ can be successfully synthesized with F₂ steady-state frequencies ("targets") that *are* independent of their vocalic contexts (Lisker, 1957). These happen to be frequencies whose closest matches in natural English speech are the values commonly reported for /ui/ (Lehiste, 1964).

There is a further similarity between English /wj/ and /ui/: just as there is no overlap in the F₂ formant frequency ranges reported for /u/ and /i/ (Peterson & Barney, 1952), so too is there a wide separation between the F₂ target frequencies for the two glides. This is of course not at all unexpected, since it would seem from the literature that the articulatory differences between the glides are as great, or even greater, than those between the vowels. This /w/-/j/ difference should hold even if we compare the two glides across different vocalic contexts. If, for example, the /w/ in /iwi/ requires tongue backing and lip rounding movements, while the /j/ in /uju/ calls for tongue fronting and lip unrounding, then these /w/ and /j/ allophones should show quite distinct F₂ frequency ranges. F₂ frequency determinations of /wj/ over a wide range of vowel contexts indicate that this is so (Lehiste, 1964), as do the results of several experiments in synthesis (Lisker, 1957; O'Connor et al., 1957).

Because spoken /w/ and /j/ are so different acoustically, reflecting their very different tongue and lip positions, it might be supposed that glides in selected contexts, as in the sequences /iwi/ and /uju/, are also necessarily disparate, both acoustically and articulatorily. However, the perceived difference between /w/ in /i/-/i/ and /j/ in /u/-/u/ need not crucially depend on such large acoustic and articulatory differences. Thus we can imagine results of a very simple operation using an old-fashioned but readily available piece of equipment—the experimenter's own vocal tract—which would make a straightforward account

of /wʃ/ perception problematical. The sequence /iwi/ is usually described as one in which the tongue moves from front to back and front again, while the lips shift from unrounded position to rounded and back to unrounded. In the sequence /uju/, on the other hand, the tongue first moves from back to front and then back again, while the initially rounded lips unround and then round again. Thus we have:

	i	w	i	u	j	u
tongue	front	back	front	back	front	back
lips	-rnd	+rnd	-rnd	+rnd	-rnd	+rnd

Let us now instead imagine an /i/-glide-/i/ sequence in which only the lips change position, while the tongue remains fixed. The glide target is then the vocal tract shaped as for the high front rounded vowel [y]. Let us also imagine another sequence, /u/-glide-/u/, this time one in which the tongue moves from back to front to back, while the lips maintain a rounded position. The glide target is again that of a front rounded vowel.

	i	ɥ	i	u	ɥ	u
tongue	f	r	o	n	t	
lips	-rnd	+rnd	-rnd		+ r	n d

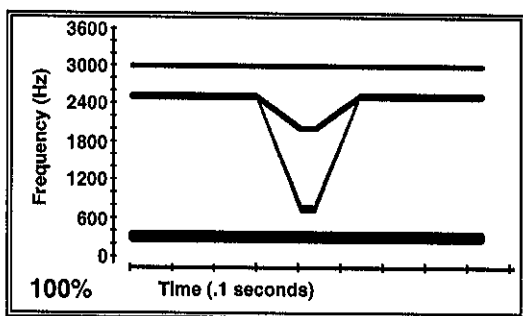
If these sequences are performed, at least one speaker-listener (the present author) is convinced that the resulting acoustic signals will be heard by English-speaking listeners as /iwi/ and /uju/.¹² If this expectation is in fact fulfilled, then we may conclude that a satisfactory /w/, at least in the /i/-/i/ environment, requires no more than a labial movement, while a phonologically adequate /uju/ can be achieved with no more than a tongue fronting maneuver.

At this point, of course, all we have is little more than a rather crude Gedanken experiment, except for the fact that two utterance types were in fact articulated and identified a number of times by at least one subject (the author). The intent was to produce [iɥi] and [uɥu], but the phonetic interpretations of the acoustic results of the articulatory movements sharing the single goal were not the same: the disyllables heard were /iwi/ and /uju/. Of course, we have no guarantee that in producing [iɥi] only the lips moved between the vowels, or that in [uɥu] only the tongue shifted position. In fact, acoustic measurements of a number of tokens of the two sequence types suggest that the effort to achieve the same target [y] (or [ɥ]) in the two contexts was a failure: the apparent F₂ target of the glide was about 500 Hz higher in one context than in the other. This failure, moreover, can hardly be explained by the phenomenon of undershoot (Lindblom and Studdert-Kennedy, 1967), since F₂ frequency at the glide midpoint was higher in the /u/-/u/ context than the /i/-/i/. If the acoustic

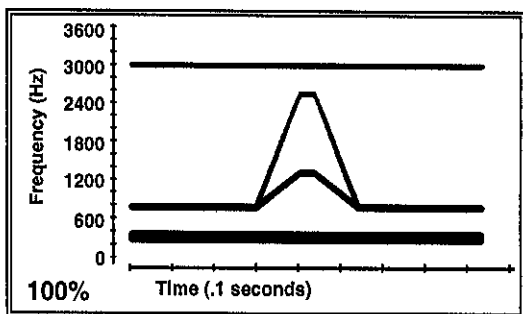
¹²In the tests in which listeners were asked whether or not they heard the palatal glide [j], the label suggested was "y."

mismatch is acceptable evidence of a failure by the speaker to perform as intended, then it appears that the glides in one or both sequences were produced with considerable *overshoot* of the single intended high-front-rounded target position. At this point, moreover, we cannot know whether the context-dependent difference in interpretation of the intended single glide is a consequence of the acoustic difference.

Since the acoustic evidence suggests that the speaker, despite his intention, did not approximate the same glide target in the two contexts, the technique of terminal analog synthesis was resorted to, in order to determine whether acoustic identity of the glide targets would elicit the same phonetic percept in the two contexts. A terminal analog software synthesizer set for series operation was used to generate two sets of patterns having the frequency and time characteristics shown in Figure 1(A and B). In each set the variables were the F₂ "target" frequency and flanking transitions: 18 acoustically different patterns in all.



/ iwi /



/ uyu /

FIGURE 1. Two sets of spectral patterns tested. In each F₁ and F₃ were fixed, while F₂ was varied as shown, so that it moved linearly to and from each of 18 "target" frequencies. In the /iwi/ set F₂ targets ranged from 750 to 2400 Hz, while for /uyu/ the F₂ targets ranged from 800 to 2500 Hz.

On the basis of preliminary trials it was thought appropriate to test the two stimulus sets separately. In one, the members of the A set were each presented five times in random order, and listeners were asked to decide whether or not they heard /iwi/, responding with either "w" or "x." In the other test the judgments elicited were either "y" (for /uju/) or "x." Thus the labeling tests did not force a choice between two competing phonemes, allowing instead for the easy "out" of an "x" response." It could therefore be expected that stimuli reported as "w" or "y" would be phonologically very convincing imitations of /iwi/ and /uju/. The responses from two linguist/phoneticians, both native American English speakers, are shown in Figure 2. Both show a broad range of F₂ frequencies judged as "w" in /i/-/i/ and as "y" in /u/-/u/. It may be noted that neither listener, in post-test discussion, reported hearing a [ɥ] glide in any of the stimuli.

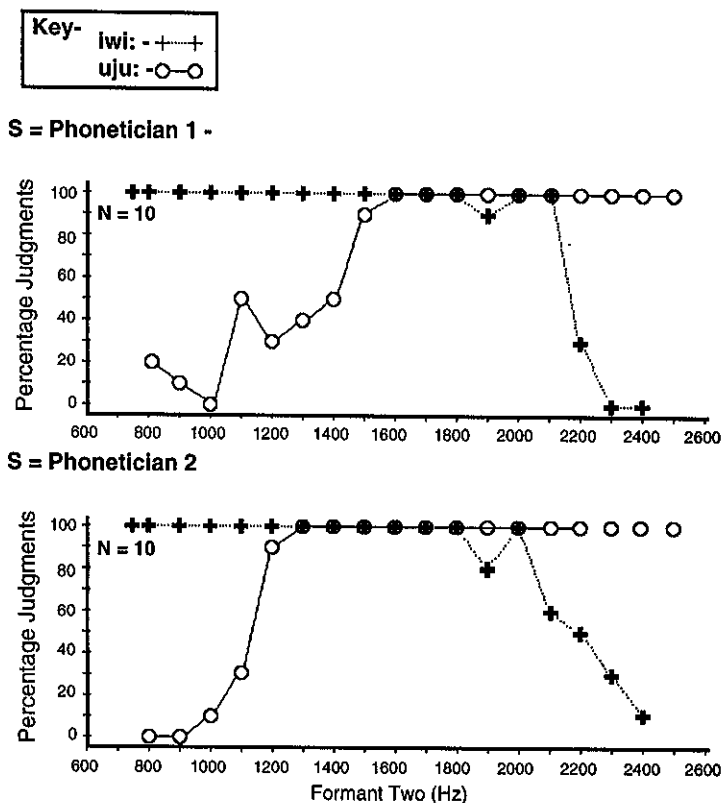


FIGURE 2. "w"/ "not-w" and "y" / "not-y" judgments by two phonetically trained listeners. Each point represents 10 independent judgments per stimulus by one listener.

These tests were also given to 20 other English-speaking subjects, none of them phoneticians, but all with some experience as subjects of speech perception experiments. Their responses (Figure 3) show a much reduced range of overlap between "w" and "y" responses, reflecting more the "noisiness" of the data than any systematic difference between their percepts and those of the two phonetically trained subjects. It may be noted that the crossover between the w and y curves falls near 80%, another indication that the two glides, phonetically and phonologically distinct as they are perceptually, do not occupy distinctly different F₂ frequency ranges.

From this experiment in synthesis it must be concluded that the failure to hear the same glide in the naturally produced (intended) sequences [iɥi] and [uɥu] is not to be explained by the speaker's failure to perform glides over the different contexts that approximated a single acoustic (primarily F₂) target. There was, to be sure, a failure by the speaker to produce glides allowing us to infer a single target, but speaker success would not have led listeners to divine the speaker's intention.

The failure to produce glides having the same acoustic target, given the speaker's intention to do just that, suggests a failure of proprioceptive judgment. Conceivably, in trying to produce [iɥi] the lips achieved a more closely rounded position than that maintained over the [uɥu] sequence. Or perhaps the tongue reached a more forward position during the glide in [uɥu] than in [iɥi]. Because of uncertainty as to the basis for the speaker's apparent failure to realize his intention, a somewhat more elaborate experiment in articulation was performed.

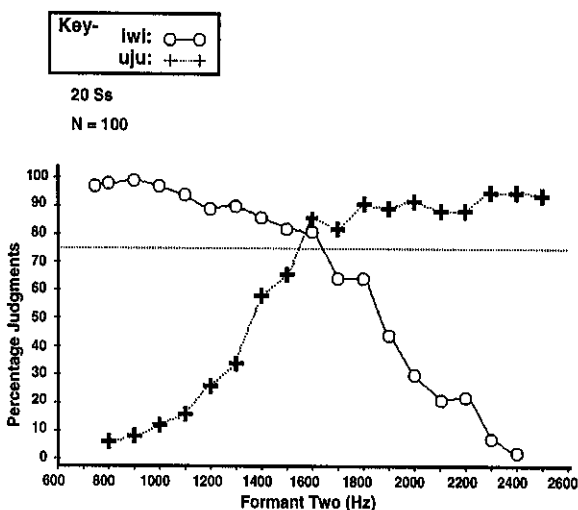


FIGURE 3. "w"/ "not-w" and "y" / "not-y" judgments by 20 phonetically untrained English-speaking listeners, each of whom made 5 independent judgments per stimulus.

Ten tokens of each of six phonetic sequences were produced by the author. Four glide targets were his goals: 1) a high back rounded shape, that of standard English /w/; 2) the high front unrounded shape of English /j/; 3) a high front rounded shape, that of French [y]; and 4) a high central rounded shape, that of IPA [ɥ]. In aiming to produce the glide [ɥ] in the context of /u/-/u/ only the tongue was consciously activated, but in /i/-/i/ the lips were very active, while the tongue moved over a restricted range.

The F₂ frequencies achieved in these articulatory exercises are given in Figure 4. In the first pair /w/ and /j/ have the widely separated F₂ frequencies we expect of those glides, where both tongue and lips are active. In the second pair, for which a single target, that of the vowel [y], was intended, F₂ frequencies for the glides were closer than for intended [wj], but differ significantly in mean values. They represent a replication of the initial attempts to achieve the sequences [iɥi uɥu], with no greater success than before. With the third pair of sequences, however, the same intended target in the different contexts, that of a high central rounded vowel [ɥ], appears to have been hit, in that there is no significant difference between the mean F₂ frequencies achieved in the two vocalic environments. However, although here we may thus have some warrant for speaking of a single glide, informally tested listeners showed no hesitation in identifying the glides in /i/-/i/ and /u/-/u/ as "w" and "y," respectively.

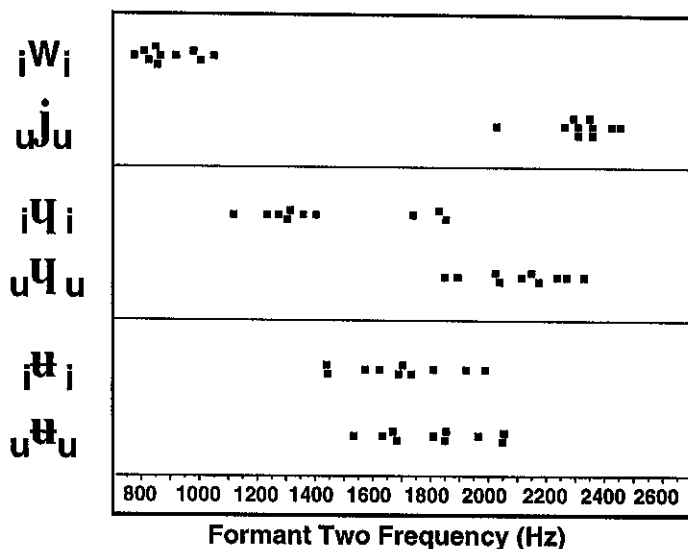


FIGURE 4. "Target" F₂ frequencies measured at the points of inflection (minimum in /i/-/i/; maximum in /u/-/u/); means of 10 productions of the (intended) phonetic strings indicated.

The extensive overlap in F_2 values found for synthesized intervocalic /w/ in /i/-/i/ and /j/ in /u/-/u/ is not consistent with the results of earlier work on glides, either in medial or in initial position, nor is it compatible with the view that /w/ and /j/ must have as targets the tongue and lip positions of the vowels /u/ and /i/. If we suppose that in the disyllables /wi/ and /uju/ there is a syllable division between the initial vowel and the glide, then we might expect about the same perceptions to hold if we delete the initial steadystate and following transition intervals of the disyllabic patterns of Figure 1. However, when the same kind of labeling tests were carried out on the truncated patterns, rather different results emerged. From Figure 5 it appears that the "w"- "y" ranges along F_2 did not overlap at all. Instead, there is a considerable gap, about 500 Hz wide, between the highest F_2 frequency heard as "w" in /wi/ and the lowest F_2 frequency that is heard as "y" in /ju/. We may note also that the F_2 ranges heard as "w" and "y" in initial position are very much narrower than those perceived as the two glides in intervocalic position.

From the results of these experiments in synthesis and deliberate articulation it appears that even if English /w/ is normally produced as a labial-velar and /j/ is an unrounded palatal approximant, at least in some intervocalic positions the labial and lingual components of the two glide articulations are not of equal perceptual weight. Thus English /wi/ can be satisfactorily produced by a lip gesture with the tongue maintaining a fixed palatal constriction, while for /uju/ the tongue fronting gesture need not be accompanied by lip movement.¹³

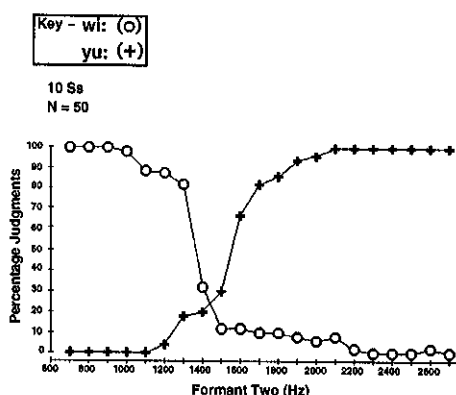


FIGURE 5. Judgments of synthesized glide-vowel patterns as "w"/ "not-w" before /i/ and "y" / "not-y" before /u/. Each point represents means of 50 judgments per stimulus, 5 by each of 10 phonetically untrained English-speaking listeners.

¹³ Its description as an "unrounded palatal semi-vowel" (Jones 1960, p. 209) surely exaggerates the role of lip position in its formation. According to Bronstein (1960) the lip position for /j/ is determined by the following vowel, being rounded before back vowels (p. 123).

This suggests that English /w/ is primarily a labial, while /j/ is a palatal approximant. Of course, if /w/ is not distinctively back and /j/ is not distinctively unrounded, then they are not to be equated phonologically with the vowels /u/ and /i/, for each of which both lip and tongue positions must presumably be specified. If the /iwi/ and /uju/ sequences are managed with the degree of gestural economy achieved in realizing them as [iɥi] and [uɥu], then the included glides should have the same target vocal tract shape, that of the front rounded vowel [y]. Why, then, do listeners presented with [iɥi]–[uɥu] seem not to know that the lip gesture in [iɥi] and the tongue gesture in [uɥu] are aimed at the same goal, inasmuch as they do not perceive the glides to be phonetically the same? After all, if the same glide *were* perceived, an explanation appealing to the single articulatory target might be readily applied. In the case of the [iɥi]–[uɥu] pair, where the gestural difference is not as stark, an explanation based on acoustic invariance would also be viable, but again only if only a single glide, not two, were reported. Since in neither of the pairs [iɥi]–[uɥu] and [iɥi]–[uɥu] was the single articulatory goal of the glides perceived, not even by the listener privy to the speaker's intention (the speaker himself), then clearly neither an invariant spectrum nor a single articulatory target serves any explanatory function in accounting for perception. Instead, in both pairs of phonetic sequences listeners perceived the different maneuvers as two distinct phonetic messages, despite the fact that they were aimed at a single articulatory target. The speaker's distinct lip and tongue movements may be construed as different ways of achieving the same goal, but the speaker's successful execution of the articulatory task, even to the point of achieving acoustic invariance, does not mean that listeners will appreciate it.

ACKNOWLEDGMENT

The support for this work provided by NIH Grant HD-01994 to Haskins Laboratories is gratefully acknowledged.

REFERENCES

- Bronstein, A. J. (1960). *The pronunciation of American English*. New York: Appleton-Century-Crofts.
- Browman C. P., & Goldstein, L. (1992). "Targetless" schwa: An articulatory analysis. In D. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 26-56). London: Cambridge University Press.
- Browman, C. P., & Goldstein, L. (in press). Dynamics and articulatory phonology. In T. van Gelder & B. Port (Eds.), *Mind as motion*. Cambridge MA: MIT Press.

- Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.
- Delattre, P. D., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 27, 769-773.
- Harris, K. S. (1974). Physiological aspects of articulatory behavior. In T. A. Sebeok (Ed.), *Current trends in linguistics* (Vol. 12, pp. 2281-2302). The Hague: Mouton.
- Harris, K. S. (1984). Coarticulation as a component in articulatory description. In R. G. Daniloff (Ed.), *Articulatory assessment and treatment issues* (pp. 147-167). San Diego, CA: College-Hill Press.
- Harris, Z. S. (1951). *Methods in structural linguistics*. Chicago: University of Chicago Press.
- Heffner, R-M. S. (1950) *General phonetics*. Madison WI: University of Wisconsin Press.
- Hockett, C. F. (1958). *A course in modern linguistics*. New York: Macmillan.
- I. P. A. (International Phonetic Association). (1949). *The principles of the International Phonetic Association*. London.
- Jones, D. (1960). *An outline of English phonetics* (9th ed.). Cambridge: Cambridge University Press.
- Ladefoged, P. (1982). *A course in phonetics* (2nd ed.). New York: Harcourt Brace Jovanovich.
- Lehiste, I. (1964). Acoustic characteristics of selected English consonants. *Publication 34 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics*. Bloomington: Indiana University Press.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- Lindblom, B. E. F., & Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America*, 42, 830-843.
- Lisker, L. (1957). Minimal cues for separating /w,r,l,y/ in intervocalic position. *Word*, 13, 256-267.
- Nearey, T. M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America*, 85, 2088-2113.
- O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. D., & Cooper, F. S. (1957). Acoustic cues for the perception of initial /wjr/ in English. *Word*, 13, 24-43.
- 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.

- Sussman, H. M., McCaffrey, H. A., & Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America*, 90, 1309-1325.
- Swadesh, M. (1947). On the analysis of English syllabics. *Language*, 23, 137-150.
- Strange, W. (1989). Evolving theories of vowel perception. *Journal of the Acoustical Society of America*, 85, 2081-2087.
- Trager, G. L., & Bloch, B. (1941). The syllabic phonemes of English. *Language*, 17, 223-246.
- Trager, G. L., & Smith, H. L., Jr. (1951). *An outline of English structure (Studies in Linguistics: Occasional Papers, 3)*. Norman: OK: Battenburg Press.