Vincent van Heuven & Annelies Hoos

Hiatus Deletion, Phonological Rule or Phonetic Coarticulation?

1. Introduction

For most, if not all, languages in the world, there is a strong preference for a regular alternation of consonants and vowels: CV–CV–CV–... In languages such as Dutch and English (and many others), however, it is quite often the case that one word ends with a vowel while the next word begins with a vowel. Dutch has two options to restore the utterance to a sequence of the canonical CV-shapes: (i) insert a glottal stop at the boundary between the two vowels, so that the two syllables remain auditorily separated by a very short consonant-like interruption or (ii) insert a semi-vowel which fluently joins the vowels across the word boundary. For unknown reasons the phonological literature on Dutch has concentrated almost exclusively on the semi-vowel insertion process, also called hiatus deletion (Booij, 1981; Trommelen & Zonneveld, 1979; Zonneveld, 1978), whereas glottal stop insertion, until recently, has received no attention (but see Jongenburger & van Heuven, 1991).

Although the phonological proposals differ in their details, the general idea is that in a sequence of two abutting vowels, a semi-vowel is inserted whose features are determined by the first vowel in the sequence. A front unrounded glide /j/ is inserted after front vowels, a back rounded glide /w/ after rounded back vowels. Opinions diverge on the feature specification of the glide that should be inserted after a rounded front vowel. This could be either /j/, /w/, or a semivowel that has no status in the phoneme system of Dutch, which would be specified by the features [-back, +round]. What makes these glide-insertion rules seem strange is that low vowels and schwa are systematically excluded: no glide can be inserted between two vowels if the first one is a low vowel or schwa. Yet one frequently observes the occurrence in Dutch of a low vowel or schwa followed by another vowel without an intervening glottal stop, i.e., two vowels smoothly joined together without an audible semi-vowel.

In earlier studies glide insertion was either implicitly or explicitly restricted to word-level phonology. In a more recent study, however, Berendsen & den Os (1987) convincingly demonstrated that the domain of glide insertion is the Intonation Phrase, i.e., the rule applies both below and above the word-level.

It is our view that the fluent linking of two abutting vowels should not be accounted for by a dedicated phonological insertion rule. We argue that the process at hand is a straightforward linking of two sounds through coarticulation. When the vocal tract moves from a configuration that is characteristic of, e.g., /i/ towards that of an open vowel /a/, a relatively slow opening gesture is executed, which generates a transition that is very much the same as, but not identical to, the one we observe during the production of a semi-vowel /j/. Similarly, the smooth transition from back rounded /u/ to /a/ passes through a trajectory that is very much alike, but not identical to, that of the semi-vowel /w/. The reason that we do not identify a semi-vowel when the transition is made the other way about, i.e., from an low vowel towards either /i/ or /u/, is that the phoneme inventory of Dutch contains no phoneme of this type. Yet, if we play a taperecording of such a closing transition backwards, we quite distinctly hear either a semi-vowel /j/ (when the target of the transition was /i/) or /w/ (when the target was /u/). To us this is convincing evidence that all these phenomena are to be accounted for by a single mechanism, viz., low-level phonetic coarticulation.

It is the purpose of the present paper to test some further consequences of our hypothesis that hiatus deletion (or: semi-vowel insertion) is not a phonological process, and need not be accounted for by a phonological rule, but is a simple case of coarticulation. We have conducted two experiments to support our view. The first experiment is an acoustic study showing that the transition sound that arises in the context of two abutting vowels is **not identical** to a proper semi-vowel /j/ or /w/. The second is a perception experiment showing that Dutch listeners do not confuse the contrast between (i) two vowels fluently joined by a "semi-vowel-like" transition and (ii) two vowels separated by an underlying semi-vowel.

2. Experiment I: Acoustic measurements of natural speech

If the transition sound that occurs when two abutting vowels are fluently joined across a word boundary is just the result of coarticulation, one would expect such a sound sequence to be different from a an otherwise identical sequence of phonemes that has a true semi-vowel at the spot of the transition sound. If a true semi-vowel would be inserted the utterance <u>Wil Marie An zien?</u> /wIl mari An zin/ ('Does Mary want to see Anne?'), it would have the same pronunciation as <u>Wil Marie Jan zien?</u> /wIl mari jAn zin/ (Does Mary want to see John?'). If, however, the two vowels of /mari <u>An</u> / are simply coarticulated we would expect this sound sequence to be shorter than the same two vowels with a true semi-vowel in between /mari <u>jAn</u>/. Secondly, in the former case one would just expect a more or less linear (or perhaps ballistic) shift of the vowel formants from /i/ to /a/, whilst in the latter case we would have to see some deviation of a straight trajectory from /i/ to /a/, suggesting the presence of an intervening sound. These diverging predictions are easily tested by systematically contrasting utterances of the above types.

2.1. Method

The basic stimulus material for this experiment consisted of short sentences that contained one word in citation form:

```
Wil MaRIE  [...] zeggen       /wIl mari  [...] zEg@/
Wil je NU  [...] zeggen       /wIl j@ ny [...] zEg@/
Wil je MOE [...] zeggen       /wIl je mu [...] zEg@/
Wil JoSEE  [...] zeggen       /wIl joze  [...] zEg@/
Wil je ZO  [...] zeggen       /wIl j@ zo [...] zEg@/
Wil je MA  [...] zeggen       /wIl j@ ma [...] zEg@/
```

The accented (capitalised) syllable ended in one of six phonologically long vowels /i,y,u,e,o,a/. The capitalised syllable attracts a contrastive accent, so that the next word will remain unaccented; Berendsen & den Os (1987) have shown that this prosodic condition is maximally conducive to glide insertion. The slots symbolised by [...] contained (quasi) minimal triads of disyllabic words with stress on the first syllable. One member of each triad began with a vowel /u,e,a/, the other two with a semi-vowel /j/ or /w/ followed by the same vowels:

```
 oe(ver)  / uv@r/        e(zel)   / ez@l/          a(ren)   / ar@/
joe(ka)   /juka/        Je(zus)   /jezOEs/        ja(ren)   /jar@/
woe(de)   /wud@/        we(zel)   /wez@l/         wa(ren)   /war@/
```

One male speaker of Dutch recorded these 54 sentences twice in random order using high quality equipment. The speaker was instructed to pronounce each sentence fluently, i.e., without any pause or interruption between any of the words.

2.2. Analysis and results

The recordings were stored on computer disk (10 KHz, 12 bits, 0,3–4,5 KHz BP). Formant frequencies and bandwidths were estimated by the split–Levinson method for robust formant analysis (Willems, 1986) using a time window of 25.6 ms that was shifted along the time axis in steps of 10 ms. Formant trajectories of the V1(G)V2–sequence were stylized (by hand) fitting each formant trajectory by maximally 5 straight line segments, as illustrated in figure 1. To this end six points were defined along the time–axis as follows:

    t1: onset of first vowel (V1)
    t2: offset of first vowel, onset of transition part
    t3: end of transition, beginning of steady state of semi–vowel (G)
    t4: end of semi–vowel steady state, beginning of second transition
    t5: end of second transition, beginning of V2 steady state
    t6: offset of second vowel (V2).

Note that the points t3 and t4 can be absent (i.e., equal to t2) when – in the case that no steady state semi–vowel occurred – the vowel transition joins V1 and V2 with monotonically increasing or decreasing functions. At each point along the time–axis the frequencies of the first (F1) and second formants (F2) were extracted. F1 is the lowest resonance generated by the vocal tract, and roughly reflects vowel height. F2 is the second lowest resonance, and corresponds roughly (and inversely) to the articulatory parameter of backness.
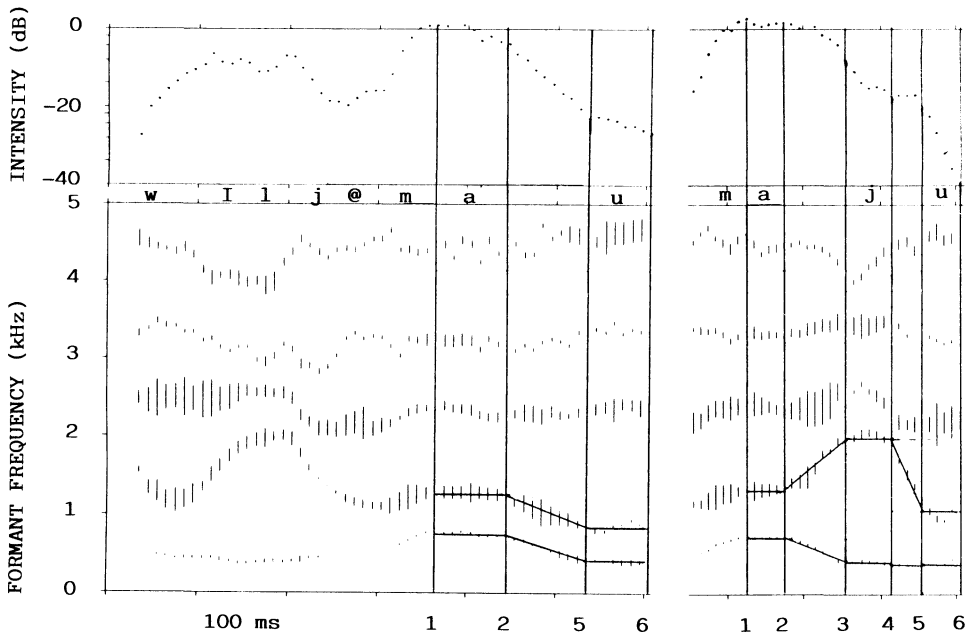
Figure 1: Segmentation points and spectral parameters for crucial V1(G)V2–sequences as determined from resograms of natural utterances of <u>Wil je MA oever zeggen</u> (left) and <u>Wil je MA joeka zeggen</u> (right).

In table 1 we present the duration of the entire V1(G)V2–sequence (the time interval between t1 and t6) as well as the duration of only the transition portion (the time interval between t2 and t5) broken down by the three types of intervening sound: /j/, /w/, or nothing (∅).

Table 1: Duration (in ms) of entire V1(glide)V2-sequence (top row) and of intervocalic transition (bottom row) broken down by type of intervocalic segment: Ø (none), /j/ or /w/. Each cell mean is based on 36 utterances.

transition segment:	Ø	/j/	/w/
entire sequence	278	330	326
transition part	80	168	147

It is quite obvious from this table that the V1GV2-sequence is some 40 to 50 ms longer when G is a true, i.e., underlying, semi-vowel /j/ or /w/ than when G is absent, in which case the transition sound could be just the reflection of coarticulation. The effect is significant by a classical one-way analysis of variance, F(2,105)=19.9 (p<.001). Moreover, the sequences including underlying /j/ or /w/ do not differ from one another, but both differ from sequences without an underlying semi-vowel (Newman-Keuls post hoc analysis of contrasts, p<.05 criterion).

We notice further that the difference is most clear when we concentrate on just the transition portion within the sequence. Here the transitions associated with true semi-vowels are about 75 ms longer than those that arise from simple coarticulation, F(2,105)=171.1 (p<.001). For this parameter, all three types of transition sound, whether /j/, /w/ or no semi-vowel, differ significantly from each other.

One may object that the above breakdown is unfair. When the linguistic prediction is that after a front vowel (/i/ or /e/) a semi-vowel /j/ will be inserted, the responses for front vowels followed by no underlying semi-vowel should only be compared with front vowels followed by underlying /j/, not with front vowels followed by underlying /w/. In order to meet this possible objection, a selection of the complete data set is now presented in table 2, such that for front vowels only underlying semi-vowel /j/ is permitted, and for back vowels only /w/. Low or central vowels have been left out of the comparison altogether.

In spite of the fact that a cleaner comparison has been made here, there is no difference between the results in table 1 and table 2. Again, total duration of the V1GV2-sequence is about 50 ms longer for true, underlying, semi-vowels than for "inserted" transition sounds, F(2,45)=12.0 (p=.001). Similarly, the transition duration is about 75 ms longer for true semi-vowels than for "inserted" transition sounds, F(2,45)=91.6 (p<.001). Both true semi-vowels do not differ from one another, neither in terms of total duration of the sequence nor in terms of the transition duration, but both types of semi-vowel differ from the condition where a transition sound is claimed to be "inserted".

In addition to differences in temporal structure between vowel sequences separated by a true semi-vowel versus those linked by a coarticulatory transition, we have looked for differences in the spectral domain. Conceivably, the vocal tract is constricted more severely for an intervening true semi-vowel (a consonant after all) than in the case of just a vocalic transition sound. Such a difference should become manifest in the spectral distance covered by the trajectories of either or both the first and second formants. In order to explore this possibility we established the maximum frequency distance simultaneously covered by F1 and F2 (i.e., by computing the Euclidean distance between two points in the F1/F2-plane) between any two of the six segmentation points defined in figure 1. In order to secure optimal comparability of distan-

ces covered in different frequency regions, all formant frequency measurements were transformed to a Bark scale (cf. Bladon & Lindblom, 1981; van Heuven, 1988) so as to reflect the sensitivity of the human hearing mechanism to differences in frequency. For practical purposes a frequency difference of 1 Bark is roughly equivalent to a third octave. The results are as in table 3.

Table 2: Selection of data from table 1 (further see text). Duration (in ms) of entire V1(glide)V2-sequence (top panel) and of intervocalic transition (bottom panel) broken down by V1 (/i,e,u,o/) and type of intervocalic segment: Ø (none), /j/ or /w/. Right-most column presents duration difference between transition segment types. Each cell mean is based on 16 utterances.

transition segment:	Ø	/j/	/w/	Δ
entire sequence V1=/i/	253	292		39
/e/	285	332		47
/u/	277		323	46
/o/	292		362	70
mean difference				51
transition part V1=/i/	78	145		67
/e/	82	168		86
/u/	88		142	54
/o/	77		167	90
mean difference				74

Table 3: Largest spectral distance (in Barks, see text) covered anywhere in V1(glide)V2 sequence, broken down by type of intervocalic segment: Ø (none), /j/ or /w/. Each cell mean is based on 36 utterances.

transition segment:	Ø	/j/	/w/
	3.3	4.0	3.6

In this table there is a small effect due to type of intervening sound on the maximal spectral distance covered anywhere in the V1(G)V2-sequence, such that the distance covered in the case of a true semi-vowel is .3 Bark (for underlying /j/) or .7 Bark (for underlying /w/) larger than when no semi-vowel is present, $F(2,105)=3.6$ (p=.031). Underlying /w/ and /j/ do not differ from one another, and only /w/ differs from the condition where no underlying semi-vowel is present.

Table 4: Selection of data from table 1 (further see text). Largest spectral distance (in Barks, see text) covered anywhere in V1(glide)V2 sequence broken down by V1 (/i,e,u,o) and type of intervocalic segment: Ø (none), /j/ or /w/. Right—most column presents difference between transition segment types. Each cell mean is based on 16 utterances.

transition segment:	Ø	/j/	/w/	Δ
V1=/i/	3.4	2.8		−0.6
/e/	2.9	3.0		−0.1
/u/	4.3		4.1	−0.2
/o/	4.0		3.9	0.1
mean difference				−0.2

When, however, we make the same selection from the data in table 3 as was done earlier in table 2, so as to ensure optimal comparability between the coarticulatory transition sound and the closest semi—vowel, no effect remains, F(2,45)<1. (cf table 4).

2.3. Conclusion

We conclude from this production experiment that there are clear and reliable differences between vowel sequences that contain a true, underlying semi—vowel (/j/ after front vowels; /w/ after back vowels) as opposed to such sequences fluently joined without an underlying semi—vowel. The differences are manifest in the temporal domain: the presence of a true semi—vowel leads to a considerably longer duration of the vowel sequence. No reliable differences were found in the spectral domain: the closing—and—opening gesture executed in between the two vowels covers roughly the same spectral distance, irrespective of the nature of the consonantal element separating the two vowels.

3. Experiment II: Perception of synthetic speech

If the transition sound that arises when two vowels are joined across a word (or morpheme) boundary were identical to a true consonant /j/ or /w/, sound sequences such as [wIlmarijar@zEg@] should be ambiguous to the Dutch listener. The [j] is either the onset of the word jaren 'years' or it is the result of a semi—vowel insertion rule that sticks in a /j/ after a non-low front vowel. If, however, ambiguity does not arise in such sequences, we may safely assume that there is a perceptual difference between a true (underlying) semi—vowel and a transition sound. Moreover, if the transition sound would just be the result of a simple coarticulation process joining the two abutting vowels, simple smoothing of formants across the segment boundary (by linear interpolation) between the two vowels should yield a convincing and acceptable sequence of two vowels, that will not be perceptually confused with a sequence of two vowels separated by a semi—vowel. This latter sequence will be longer and contains a consonantal sound segment between the two vowels, which may give rise to a larger spectral trajectory.

We decided to test the potential perceptual ambiguity of these two types of sound sequences (VV versus VGV) using synthetic rather than human speech. Only by speech synthesis can we ensure that the two vowels are joined by linear interpolation and that the two types of sound sequence do indeed contain the intended differences in segment structure. Moreover, since the transitions between adjacent sound segments should be the result of simple linear smoothing, we could not use "standard" synthesis schemes for Dutch such as diphone synthesis (which contains humanly produced transitions between adjacent sounds)

or allophone synthesis (which models coarticulation processes in a more complex manner). Instead we adopted the method of concatenating parametrised sound segments that have been excerpted from coarticulatory neutral contexts (which we have called "neutrones", hence "neutrone synthesis", cf. van Bezooijen, 1990).

3.1. Method

A subset of the sentences used in the production experiment were generated by the neutrone synthesis program developed in our laboratory (Guijt, 1989). The following 42 V1(G)V2-combinations were selected:

V1	G	V2	V1	G	V2
i	Ø,j	a,e,u	e	Ø,j	a,e,u
y	Ø,j,w	a,e,u	o	Ø, w	a,e,u
u	Ø, w	a,e,u	a	Ø,j,w	a,e,u

The necessary sound segments were concatenated, and given an appropriate intonation contour with standard declination and a standard 6 semitone rise-fall accent on the syllable containing V1. Utterances were truncated after V2, so that the only potential cue to differentiate between, e.g., Wil Marie oever/joeka zeggen? [wIl mari uv@r/juka zEg@] will be in the transition between V1 and V2. All the relevant synthesis parameters, i.e., formant frequencies, bandwidths and intensity, of the remaining part of the utterances were then smoothed by linear interpolation over a 50 ms time window. Figure 2 illustrates the spectral make-up of the crucial portion of the utterances containing [iØa, ija, iwa].
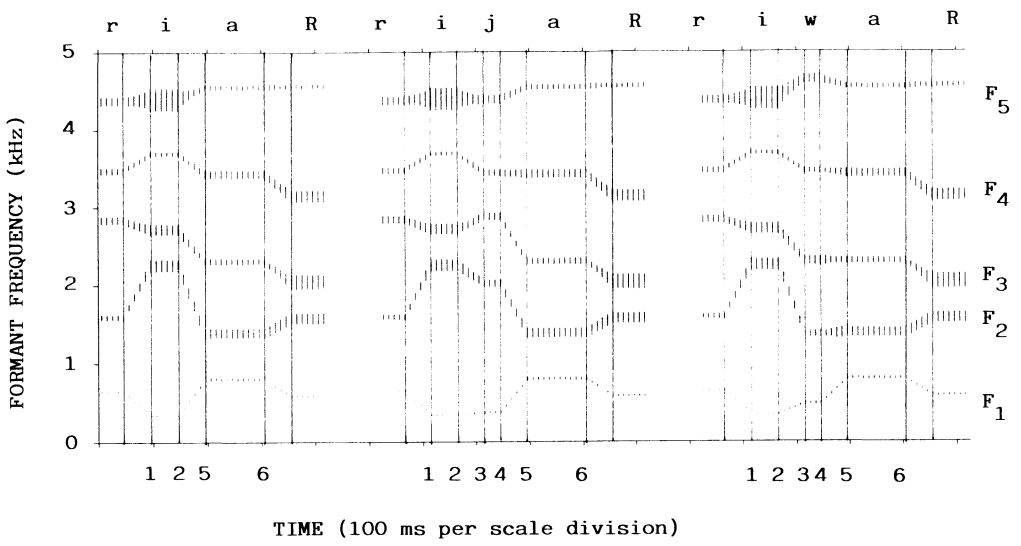

Figure 2: Spectral trajectories of F1 through F5 and B1 through B5 for concatenated neutrones synthesizing [..riØar.., ..rijar.., ..riwar..] as in Wil Marie aren/jaren/waren zeggen.

The 42 stimuli were recorded onto audio tape in two different random orders with 10s interstimulus intervals (onset to onset).

Seventeen Dutch listeners were given answer sheets that contained, for each stimulus, three printed response alternatives, one with a target word beginning with j, one with w and one with a vowel. For instance, if the stimulus was Wil Marie oe [wIl mari‾u], the response alternatives were Wil Marie joeka/woede/ oever zeggen. Subjects were instructed to indicate for each response alternative along a scale from 1 to 10 how acceptable the alternative would be given the audible stimulus (where 1 stood for "very unacceptable" and 10 for "very acceptable"). If a subject marked the three response alternatives joeka/woede/ oever with 8, 1 and 4, respectively, this should be interpreted as follows: the audible sound sequence is an acceptable token of the beginning of joeka, totally unacceptable if the intended word were woede, and somewhat less unacceptable if the intended word were oever. If the transition sound between Marie and oever is identical (in the listener's conception of the Dutch phonological system) to a semi-vowel /j/, the response alternatives joeka and oever should receive equal (high) acceptability scores, whilst woede should be an unacceptable interpretation of the stimulus. If, on the other hand, the fluent transition between two vowels is different, in the native listener's conception of Dutch, from a semi-vowel, this should be reflected in the acceptability ratings: oever would then be a more acceptable interpretation of the stimulus than joeka.

3.2. Results

Table 5 presents the mean acceptability rating for each reponse alternative for or each of the 42 stimuli. Cell means are based on 34 responses each.

When the stimulus contains the two vowels V1 and V2 simply joined by linear interpolation (leftmost column in table 5), those response alternatives are given the highest acceptability ratings that have target words beginning with a vowel (mean rating 6.7). When V1 is a front vowel (/i,e/) the alternative with /j/ is more acceptable than the alternative with /w/. When V1 is a back vowel (/u,o/) the alternatives with /w/ are preferred. When V1 is /y/ (front rounded) the alternative with /w/ is deemed slightly more acceptable than the one with /j/. Clearly, the sound that arises as a consequence of joining two vowels should be /j/-like rather that /w/ after front vowels and /w/-like rather than /j/-like after rounded (rather than back) vowels. In all cases, however, the transition sound is less acceptable as a token of a semi-vowel than as a token of a coarticulatory transition sound. On average the interpretation as a V1V2-sequence is 1.4 points more acceptable than the interpretation as a sequence of two vowels separated by a semi-vowel (either /j/ or /w/, whichever yields the most acceptable reading), $t(547)=9.2$ (p<.001).

If the stimulus contains an intervocalic glide /j/ (middle column of table 5) the response alternatives with /j/ receive the highest acceptability ratings: 6.6 on average. The mean acceptability rating of the second most acceptable alternatives is 1.8 points less, $t(358)=7.0$ (p<.001).

When an intervocalic /w/ was generated in between V1 and V2 (right hand column in table 5), response alternatives with /w/ are rated as the most acceptable alternative: 6.5. The second most acceptable alternative scores 1.6 points less, $t(367)=8.7$ (p<.001).

The results of the perception experiment suggest that Dutch listeners expect the phonetic structure of an underlying V1V2-sequence to be different from that of a V1GV2-sequence, i.e., containing an underlying semi-vowel /j/ or /w/ separating the two vowels. We conclude that our listeners apply different perceptual norms to V1V2-sequences with and without an intervocalic underlying semi-vowel. Stated more bluntly, listeners do not want to hear a semi-vowel when is is absent on the underlying level, and they want to hear one when it is underlyingly present. It is not the case that vowel-vowel sequences with and without an intervocalic semi-vowel are perceptually identical.

Table 5: Mean acceptability rating of 42 V1(G)V2–stimuli broken down by V1–V2 combination (vertically), transition sound G (horizontally) and by response alternative (horizontally). Cell means are based on 34 responses each. When the response alternative is identical to the stimulus, the rating is indicated in bold face. The second most acceptable alternative is underlined.

stimulus	C = ∅			C = j			C = w		
V1–V2\C	∅	j	w	∅	j	w	∅	j	w
i–a	**6.9**	<u>5.4</u>	4.1	<u>6.0</u>	**5.1**	4.3			
i–e	**7.2**	<u>6.1</u>	3.6	<u>5.2</u>	**5.3**	4.6			
i–u	**7.0**	<u>5.3</u>	4.0	<u>5.3</u>	**6.3**	4.5			
e–a	**6.6**	<u>6.3</u>	3.5	<u>4.6</u>	**7.0**	4.1			
e–e	**7.2**	<u>5.2</u>	3.6	<u>4.8</u>	**6.6**	4.1			
e–u	**6.3**	<u>5.9</u>	3.7	<u>4.8</u>	**7.3**	4.4			
u–a	**7.0**	3.3	<u>5.2</u>				<u>5.5</u>	4.4	**5.1**
u–e	**6.4**	4.5	<u>4.7</u>				<u>5.9</u>	3.4	**6.1**
u–u	**6.4**	3.6	<u>4.9</u>				<u>5.4</u>	3.4	**6.0**
o–a	**6.4**	3.8	<u>6.1</u>				<u>5.0</u>	3.8	**5.0**
o–u	**6.3**	4.4	<u>5.6</u>				<u>4.9</u>	3.5	**7.1**
o–u	**6.2**	3.5	<u>5.8</u>				<u>4.7</u>	4.0	**6.4**
y–a	**7.0**	4.3	<u>4.4</u>	<u>4.8</u>	**6.6**	3.8	5.1	<u>5.1</u>	**6.6**
y–e	**6.5**	3.3	<u>5.5</u>	<u>4.9</u>	**6.6**	4.0	<u>4.4</u>	4.0	**7.2**
y–u	**6.6**	4.6	<u>4.8</u>	<u>4.6</u>	**7.2**	3.9	<u>5.3</u>	3.9	**6.8**
a–a	**7.1**	3.8	<u>4.0</u>	4.0	**6.7**	<u>4.5</u>	<u>3.9</u>	3.3	**7.1**
a–e	**6.1**	3.9	<u>5.7</u>	3.8	**7.1**	<u>4.0</u>	<u>4.1</u>	3.4	**7.8**
a–u	**7.0**	3.3	<u>4.9</u>	4.6	**7.2**	<u>4.3</u>	<u>4.9</u>	3.8	**6.6**
mean	6.7	<u>5.3</u>		<u>4.8</u>	6.6		<u>4.9</u>		6.5

4. General conclusion and discussion

The results of both experiments converge. The acoustic measurements performed on tokens of V1V2-sequences with and without an underlying semi–vowel reveal clear and systematic differences in temporal organisation between the two types. When the sequence contains an underlying (true) semi–vowel is lasts significantly longer than when two vowels are joined across a word boundary. The transition sound that links the two vowels in this case is therefore different from a semi–vowel.

The results of the perception experiment indicate that Dutch listeners know that the two kinds of V1V2 sequences should be different. Listeners know that for a V1V2-sequence with an intervocalic glide to be acceptable, it should contain an audible semi–vowel segment, that should not be there when the utterance is to be an acceptable V1V2-sequence without an underlying semi-vowel.

The results of these two experiments force us to conclude that there is no evidence for a glide-insertion rule in the phonology of Dutch, at least not one that operates across word boundaries. It should be pointed out that the results of our experiments do not exclude the possibility of a glide insertion rule that operates within words.

Simply joining two adjacent vowels across a word boundary by linear inter-polation of their spectral parameters, without changing the duration of the sequence, is systematically rated more acceptable than joining the two vowels by inserting a semi-vowel. Consequently, low-level phonetic coarticulation provides a simpler and more plausible account of hiatus deletion across word boundaries than does semi-vowel insertion. We would propose accordingly that the semi-vowel insertion rule be eliminated from the phonology of Dutch. Rather than researching glide-insertion, our attention should be focussed on the question when vowel-vowel sequences will be broken up by glottal stop insert-ion, and when they are fluently joined. We suggest that there should be a phonological rule for glottal stop insertion (cf. Jongenburger & van Heuven, 1991) that applies under restricted conditions; if the rule does not apply, the default is that vowels are smoothly joined. Vowel-onto-vowel coarticulation generates a transition sound that often bears a certain resemblance to a semi-vowel. This will be the case when V1 is a non-low vowel, or in more phonetic terms, when the offset of V1 has a relatively low F1 frequency. When V1 is a low vowel, or when its offset is characterised by a relatively high F1 frequen-cy, the resulting glide does not resemble a semi-vowel. We have learned from the present study that the coarticulatory transition sound is nor should be identical to a semi-vowel.

References

BLADON, R.A.W., LINDBLOM, B.E.F.
1981   Modeling the judgment of vowel quality differences, Journal of the Acoustical society of America, 69, 1414-1422.

BERENDSEN, E., OS, E. DEN
1987   Glide insertion: domains, speech rate and phonetic prominence, in F. Beukema, P. Coopmans (eds.): Linguistics in the Netherlands 1987, Dordrecht: Foris, 13-20.

BEZOOIJEN, R. VAN
1990   Evaluation of speech synthesis for Dutch: comparison of synthesis systems, intelligibility tests and scaling methods, ASSP-report no. 22, Stichting Spraaktechnologie, Utrecht.

BOOIJ, G.E.
1981   Generatieve fonologie van het Nederlands [Generative phonology of Dutch], Spectrum, Utrecht.

GUIJT, T.P.
1989   Automatische tekst naar spraak omzetting op basis van neutroonsynthese [automatic text to speech conversion using neutrone synthesis], Afstudeerverslag Hogere Informatica Opleiding, Haagse Hogeschool.

HEUVEN, V.J. VAN
1988   De waarneming van spraak [The perception of speech], in M.P.R. van den Broecke (ed.): Ter Sprake, spraak als betekenisvol geluid in 36 thematische hoofdstukken, Foris, Dordrecht, 73-103.

JONGENBURGER, W., HEUVEN, V.J. VAN
1991   The distribution of (word initial) glottal stop in Dutch, this issue.

TROMMELEN, M., ZONNEVELD, W.
1979   Inleiding in de generatieve fonologie [Introduction to generative phonology], Coutinho, Muiderberg.

WILLEMS, L.F.
1986   Robust formant analysis, IPO Annual Progress Report, 21, 34-40.

ZONNEVELD, W.
1978   A formal theory of exceptions in generative phonology, Peter de Ridder Press, Lisse.