2024. *Proc Ling Soc Amer* 9(1). 5709. https://doi.org/10.3765/plsa.v9i1.5709.

Realizations of [j] vs. hiatus in different vocalic contexts

Anya Hogoboom*

Abstract. Glide-like formants can arise either through articulation of a glide or through articulation of hiatus, which often results in similar formant movements (e.g. *she yachts*, *she ought*). Davidson & Erker (2014) established that the glide-like formant movements are measurably different from actual, phonological glides. The current study compares a wider range of vowel environments to investigate the different realizations of [j]. Analysis of modal V#(j)V sequences finds significant differences in intensity change and duration in most vowel environments examined, and additionally finds significant differences in formant movements for one ambiguous vowel environment. However, the large degree of overlap in acoustic properties is noted, casting doubt that the differences could be strongly disambiguating perceptually. Additionally, this work gives a detailed picture of the multitude of different ways in which a phonological [j] can be acoustically realized.

Keywords. glides; hiatus; formant transitions; English

1. Introduction. Glides typically have formant structures not unlike vowels, albeit characterized by constant change. Glide-like formants can arise either through articulation of a glide or through articulation of two adjacent vowels of different heights, which result in similar formant movements. The glide [j] is typically characterized by a rise in F1 and a fall in F2, formant movements which also occur in the articulation of a hiatus environment of a high front vowel followed by a mid central vowel, as shown in Figure 1.

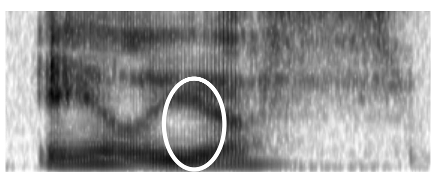

Figure 1. Spectrogram of *alias* [eɪ.li.əs], with the formant movements between [i] and [ə] circled

These glide-like formant movements also occur when hiatus arises at word boundaries, for example, in the phrase *ugly oak* [ʌgli oʊk]. Looking at such cases, Davidson & Erker (2014) established that these glide-like formant movements are measurably different from actual, phonologically-present glides. They found that V#jV sequences are longer, have a greater intensity difference between either vowel and the intensity minimum of the sequence, and have a significantly different F2, compared to V#V, at the intensity minimum of the sequence.

The current study compares a range of vowel environments to investigate the different realizations of [j] and seeks to clarify how robustly-present the differences between V#jV and V#V sequences are. V#jV sequences are referred to as such and the [j] within them as a *phonological*

* I am grateful for the support of this work through a Faculty Research Grant from William & Mary. I would like to thank the undergraduate research assistants with whom I designed and carried out this study: Abram Clear, Abby Fergus, and Elizabeth Maneval. Thanks are also due to my current research assistants for helpful discussion of these results: Carolyn Bonafede and Naomi Utgaard. Thank you also to participants of William & Mary's Linguistics Research group, particularly Kate Harrigan, and audience members at the 100th Annual Meeting of the LSA for helpful feedback, comments, and suggestions. Author: Anya Hogoboom, William & Mary (ahogoboom@wm.edu)

[j]. V#V sequences are either referred to as such, or as *hiatus*, and the formant transitions within them, when glide-like, are referred to as an *acoustic [j]*.[1]

Davidson & Erker examined the acoustic glides that can occur in two environments: the classic [+high]#[-high] hiatus discussed above and the near steady-state that results from two [+high] vowels in sequence, but did not break out differences between the two. The (near) steady-state glide that results when preceded and followed by a high vowel is noteworthy for its lack of formant movement. While glides canonically start from a high position in the vowel space and move to a lower one, they can also be articulated in reverse, from a low to a high position and occur so naturally if the preceding vowel is lower than the following one. These three environments can result in ambiguous V#(j)V sequences, as the vowel formant movements are extremely similar to those of an actual [j]. While the study described below was set up to investigate these three ambiguous environments, it was found that several phrases patterned like a fourth type of ambiguous V#(j)V sequence, one that is similar to a classic [j] but with a much shorter fall. Thus the study reports results regarding four vocalic environments which can be articulated such that they are ambiguous between the presence and absence of a glide. The work here also looks at an unambiguous environment; namely, where both vowels are [-high] and so when there is a phonological glide the speaker must specifically raise their tongue position to produce one.

2. Production study.

2.1. Participants. Forty-seven native speakers of American English (aged 18-22, mean age 19; 42 females, 5 males) participated in the study, all of whom were undergraduates at William & Mary and received participant pool credit.

2.2. Study design. Two-word phrases that created (near) minimal pairs of V#jV and V#V were created for each of the initially-targeted four vocalic environments. Subsequently, it was found that one pair that was meant to be a "reverse" [j] and two pairs that were meant to be (near) steady-state [j]s actually behaved as a further type of ambiguous V#(j)V sequence, and so these will be broken into their own category here for consistency. This results in five vocalic environments investigated, as laid out in Table 1.

Environment	Description	Name	Ambiguous?	Example
1	Between non-high vowels		No	[ɑ#(j)ɑ]
2	Between high front and non-high vowels	Classic [j]	Yes	[i#(j)ɑ]
3	Between non-high and high front vowels	Reverse [j]	Yes	[ɑ#(j)i]
4	Between high front vowels	Steady-state [j]	Yes	[i#(j)i]
5	Between high-articulated vowels with neighboring liquids or nasals	Short Classic [j]	Yes	[ri#(j)ir]

Table 1. The five vocalic environments

The initial goal was to have four pairs of phrases of each type of vocalic environment, but the numbers ended up less even because of unanticipated phrase behavior, as mentioned above. One

[1] While Davidson & Erker examined sequences that give rise to an acoustic [j] and those that give rise to an acoustic [w], the work here looks specifically at [j] for simplicity. As an acoustic [w] arises from a vowel sequence that starts with [u], and [u] is a less-commonly found vowel than [i]. Further, the formant movements with [j] (or [j]-like) articulation change more than [w]-like ones which makes them somewhat easier to examine.

pair of phrases that was initially formulated to be an example of Environment 3 (Reverse [j]) had such high articulations of what was meant to be a mid [ə] that they were reclassified as Environment 4 (Steady-state [j]). One further Environment 3 pair and two Environment 4 pairs were articulated with a small rise in F1 and small fall in F2 and so were reclassified as the newly-created Environment 5 (Short Classic [j]). Thus Environments 1 and 2 had four pairs, Environment 3 had two pairs, and Environments 4 and 5 had three pairs. A further four unpaired phrases were constructed in which the first word ended in a consonant and the second began with a phonological [j], although these are not discussed further here.

Carrier phrases in a question and answer formant were constructed for each of the two-word phrases which would serve the dual purpose of introducing the target phrase in the question portion and encourage as natural as possible production of the phrase in the answer portion with any emphasis or contrastive focus on a later part of the sentence. A pair of question and answer sentences (for a Vowel Environment 2 (Classic [j]) are shown in (1)

(1) a. Q: Is the story that she yachts on the lake every week or every month?
A: The story is that **she yachts** on the lake every month.
b. Q: Is the claim that she ought to the party or that she ought to go home?
A: The claim is that **she ought** to go home.

These Q/A pairs were randomized into PowerPoint slides in ten different orders. The complete list of stimuli answer sentences is given in the Appendix.

2.3. PROCEDURE. Participants were told they were going to read question-and-answer pairs and were asked to read each as naturally as possible, as if they were having a conversation with a friend. Recording took place in a sound-attenuated booth, where the participant wore a head-mounted Shure WH30 microphone connected to a Tascam DR-100 recorder. Participants read each question answer pair that was displayed on a single slide of a (pdf of a) PowerPoint and then progressed to the next slide (self-paced). Participants were instructed that if they stumbled or otherwise did not feel a sentence had come out well to re-read that sentence.

2.4. MEASUREMENT PROCEDURE. Four datasets were created. To create Dataset 1, the target V#(j)V sequence in each answer was coded for whether it was produced with modal or creaky voice, where “creaky” included cases with creaky voice during the formant transitions, those with a glottal stop breaking the two syllables, and those with a glottal stop and surrounding creak. Cases in which the formant transitions were produced with modal voice but creak was produced later in the second vowel were treated as modal, in line with the conclusions of Davidson & Erker who concluded that creak in V2 is unrelated to creak used to signal hiatus. The cases (N=4) with global creak were excluded, which resulted in 1,348 data points.

To create Dataset 2, the modal voice pronunciations of the target two-word phrase that were a part of Dataset 1 (N=694), and additional V#V modal voice sequences from the question portion (where the sequence in the answer was not also produced with modal voice, N=21) were delineated in Praat (Boersma & Weenink 2023). These two-word phrases were then extracted and the resulting wav files delineated by FAVE (Rosenfelder et al. 2022) and then hand-adjusted so that there was an interval from the onset of the first vowel in the V#(j)V sequence to the offset of the second. The duration and 5 RMS amplitude measurements taken at 20%, 40%, 50%, 60%, 80% of each V#(j)V sequence were extracted via Praat script (modified version of Lennes 2003).

Dataset 3 expands set on the set of phrases used in Dataset 2 to also include V#V phrases produced with creaky voice but trackable formants, based on visual inspection of Praat's formant tracking. For example, if a full glottal stop was made, formants were not trackable through the V#V sequence. But in many cases where the hiatus was produced with creaky voice, Praat was able to track the formants and these sequences were included. If the target V#V phrase was not produced with trackable formants in the answer portion, the question portion was examined and the phrase there extracted if there were trackable formants. Extracted phrases were then delineated with FAVE and hand-adjusted to capture the vowel-to-vowel interval as in Dataset 2. The duration of each sequence was extracted via the Praat script.

Dataset 4 is a subset of Dataset 2, as it consists of only modal voice V#(j)V sequences, with those sequences in which a portion of the V#(j)V sequence was too constricted to track formants (N=90) taken out, resulting in 625 data points. F1 and F2 were extracted at 20%, 40%, 50%, 60%, and 80% through each V#(j)V sequence via the Praat script.

3. Results.

3.1. TYPES OF GLIDES. Visual comparisons of the spectrograms of a sample pair from each of the five V#(j)V Vowel Environments is given, along with a vowel plot showing the individual and average F1 and F2 trajectories, taken from Dataset 3, minus V#(j)V sequences with non-trackable formants. Delineations shown for the spectrograms are approximate (only entire vowel-to-vowel sequences were analyzed so individual sound delineations are for visual purposes only) and each set of phrases shown is said by the same speaker. All vowel plots were made using the phonR package (McCloy 2016) in RStudio (RStudio 2022, R Core Team 2021).

As expected, V#jV and V#V are not ambiguous in Vowel Environment 1. When a [j] is produced, it requires a fall in F1 and rise in F2 in order to be in a position to create the classic rise in F1 and fall in F2, as can be seen the spectrogram comparison in Figure 2 and in the vowel plot is shown for all tokens and their averages in Figure 3.

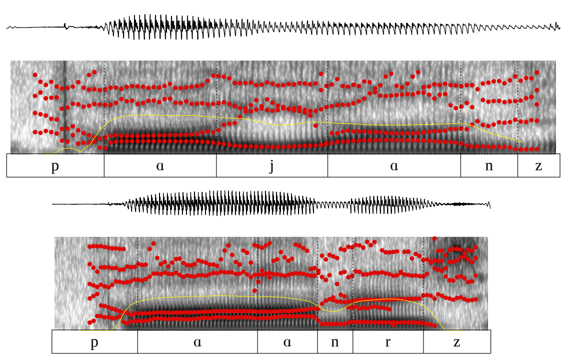

Figure 2. Spectrograms of *pa yawns* and *pa honors*

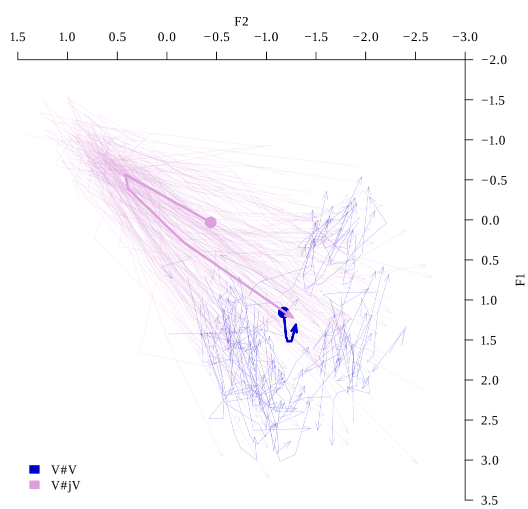

Figure 3. Trajectories of individual and average unambiguous V#(j)V sequences

It can be seen that despite having the two trajectories being from and two the same vowels, when a [j] is present it seems to nudge the first vowel to a higher, fronter starting position. When we break this down by V#(j)V pair as shown in Figure 4, we see that this occurs, unsurprisingly in the two pairs with a word-final [ə] in the first word of the phrase (bottom two in Figure 4), but also, perhaps more surprisingly, with a word-final [oʊ] in the first word of the phrase, although a word-final [ɑ] is unaffected.

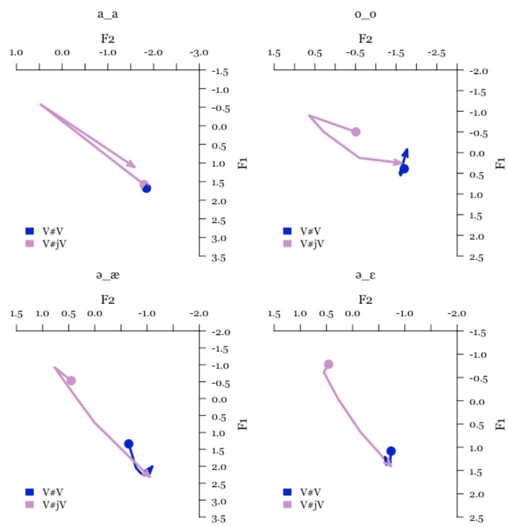

Figure 4. Average trajectories of Vowel Environment 1 V#(j)V sequences by pair

Vowel Environment 2 is the classic ambiguous case, and indeed we see a large overlap in the trajectories of the glide and the formant transitions when between a high and a non-high vowel, evident from the spectrogram comparison in Figure 5 and shown in Figure 6. On average, the V#jV sequence shows a more extreme start and end point to the surrounding vowels, compared to the V#V sequence.

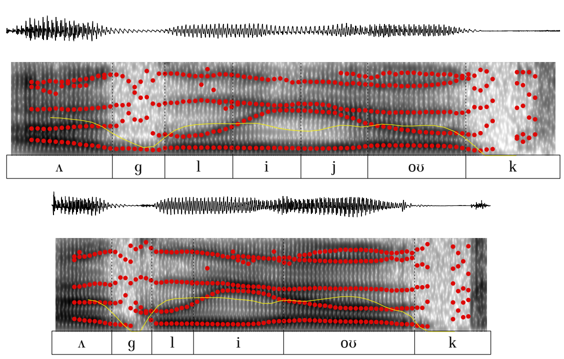

Figure 5. Spectrograms of *ugly yolk* and *ugly oak*

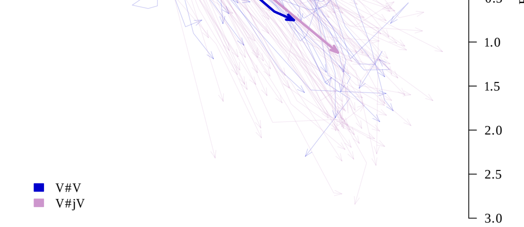

Figure 6. Trajectories of individual and average ambiguous Classic V#(j)V sequences

We again see overlap, clear in both the spectrogram in Figure 7 and the vowel plot in Figure 8, in Environment 3, the reverse [j] condition, caused by movement from a non-high vowel to a high vowel.

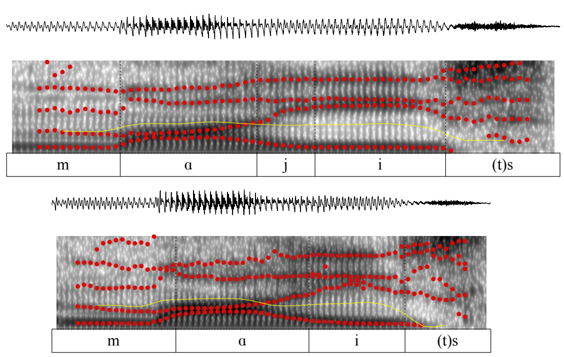

Figure 7. Spectrograms of *ma yeets* (slang for 'throws') and *ma eats*

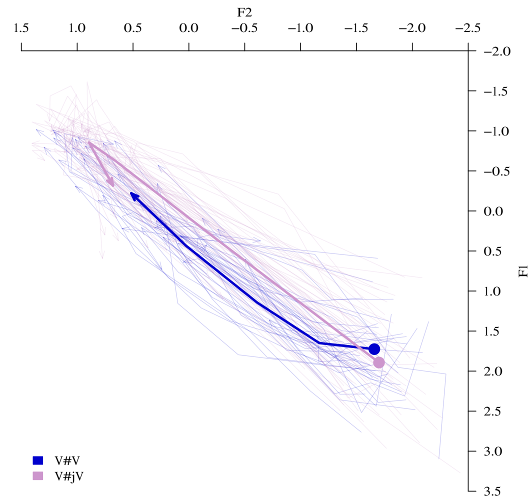

Figure 8. Trajectories of individual and average ambiguous Reverse V#(j)V sequences

Environment 4 is again ambiguous, but because we do not find movement for a [j] when between two high vowels, both V#V and V#jV sequences look like the hiatus condition. An example is shown in the pair of spectrograms in Figure 9 and the trajectories are shown in the vowel plot in Figure 10.

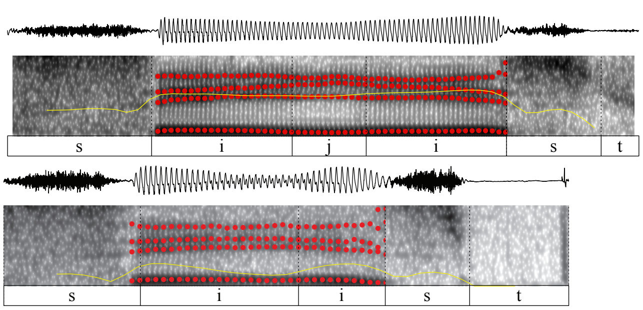

Figure 9. Spectrograms of *see yeast* and *see east*

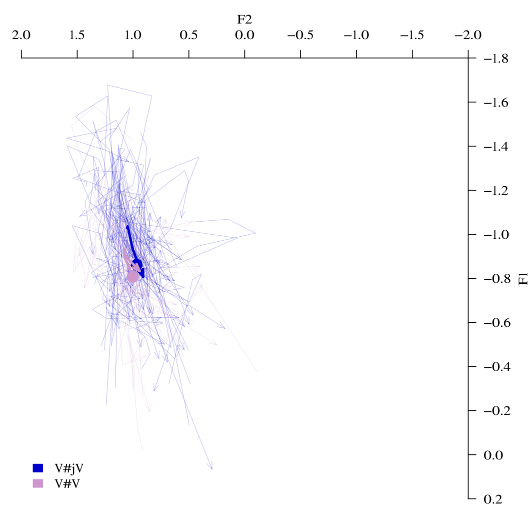

Figure 10. Trajectories of individual and average ambiguous Steady-state V#(j)V sequences

Finally, the data for Environment 5 consists of those three pairs of phrases in which a high-articulated vowel was followed by a high vowel with neighboring liquids or nasals and which resulted in a shorted version of the classic [j] trajectory, referred to as Short Classic [j]. The formant movement is not particularly evident in the spectrograms in Figure 11, but each of the three pairs were placed into this group because when mapped out in the vowel space, they showed a trajectory like the average shown in Figure 12.

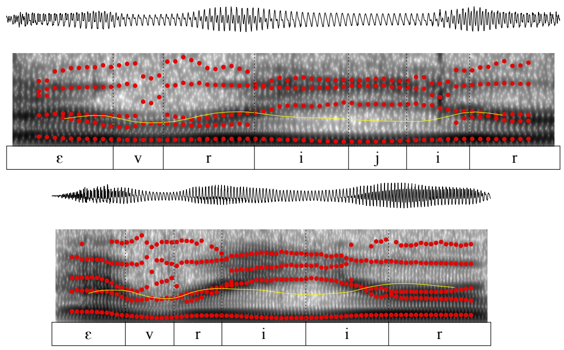

Figure 11. Spectrograms of *every year* and *every ear*

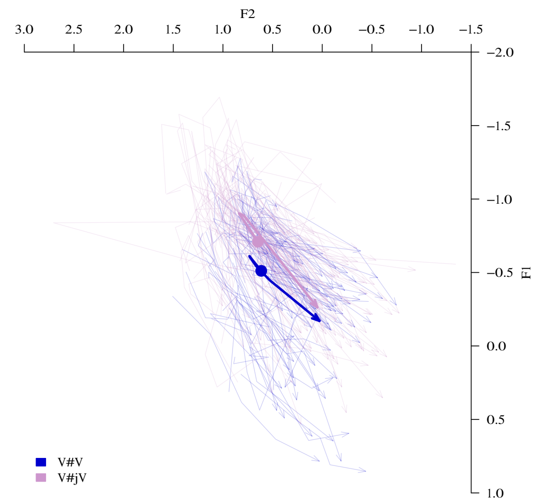

Figure 12. Trajectories of individual and average ambiguous Short Classic V#(j)V sequences

It is worth noting the many different possible trajectories that a V#jV sequence can take, which results in glides which may appear to fall in the vowel space (Classic [j] and Short Classic [j]), rise in the vowel space (Reverse [j]), rise and then fall (unambiguous type), or not involve movement (Steady-state [j]). These five different V#jV trajectories are shown in Figure 13.

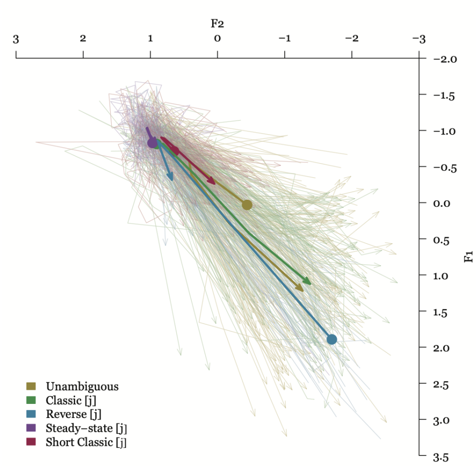

Figure 13. Comparison of V#jV trajectories

3.2. USE OF MODAL VS. CREAKY VOICE. Analysis of Dataset 1 (as described in §2.4) found that the V#V formant transitions (in the answer sentences) were overwhelmingly produced with creaky voice (625/686 =91.1% creaky), where as phonological [j]s were overwhelmingly produced with modal voice (633/662 =95.6% modal). An example of this difference between phonological [j]s and hiatus is shown in Figure 14.

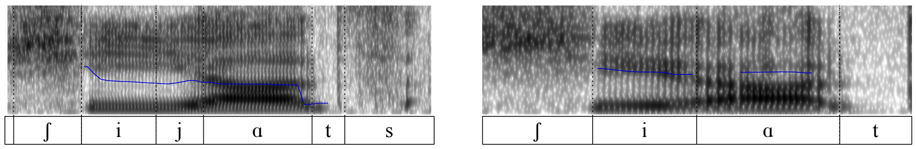

Figure 14. Spectrograms of *she yachts* with modal voice and *she ought* with creaky voice

A binomial logistic regression was run in SPSS on dependent variable *Voice* (two levels: Modal, Creaky), with independent variables *Vowel_Environment* (five levels: Unambiguous, Classic [j], Reverse [j], Steady-state [j], Short Classic [j]), *V#(j)V* (two levels: [j], hiatus), the interaction term, and with *Subject* as a blocking factor. The modal is given in Table 2.

	Wald ChiSquare	*df*	Sig
(Intercept)	3.648	1	0.056
Vowel_Environment	8.322	4	0.080
V#(j)V	258.854	1	<0.001
Vowel_Environment*V#(j)V	15.290	4	0.004
Subject	85.408	46	<0.001

Table 2. Binomial logistic regression with DV=*Voice*

Pairwise comparisons found hiatus in Vowel Environment 2 (Classic [j]) to be significantly less likely to be produced with creak than hiatus in Vowel Environments 1 (unambiguous), 3 (Reverse [j]), or 5 (Short Classic [j]) ($p \leq 0.035$).

The fact that creaky voice is equally or more likely to be used in the unambiguous V#V Vowel Environment than in the four ambiguous ones means that speakers are not using creaky voice specifically to disambiguate V#V sequences from V#jV sequences.

3.3. INTENSITY. Dataset 2 (as described in §2.4) took the RMS amplitude from modal voice V#(j)V sequences, measuring at 20%, 40%, 50%, 60%, and 80% of each sequence's duration. In the following analysis, the amplitude of the first vowel in the sequence is taken to be the measurement taken at 20%, the amplitude the second vowel is taken to be the measurement at 80%, and the phonological or acoustic [j] is taken to be the amplitude minimum of the three measurements in-between those two.

The intensity change from the first vowel (V1) to the phonological or acoustic [j] was calculated as $Amplitude\_V1 - Amplitude\_[j]$, and the intensity change from the second vowel (V2) to the phonological or acoustic [j] was calculated as $Amplitude\_V2 - Amplitude\_[j]$. Given that consonants are generally more constricted than vowels we would expect a the intensity difference between either vowel and the intensity minimum to be greater in the V#jV condition than in the V#V condition. Figure 15 shows the intensity difference from the first vowel, with the number of tokens of each noted along the top of the box plot. The number of hiatus condition modal voice tokens is very low in some cases, like in Vowel Environment 3 (Reverse [j]) but their intensity difference distribution looks very similar to that found in Vowel Environment 2 (Classic [j]) where there are more tokens.

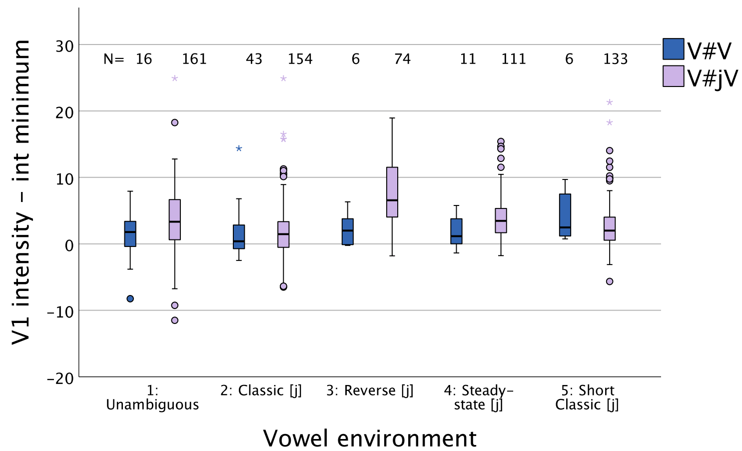

Figure 15. Intensity difference between between V1 and the intensity minimum

A Generalized Linear Model (GLiM) was run in SPSS with the dependent variable *Intensity_change_from_V1*, with independent variables *Vowel_Environment*, *V#(j)V*, their interaction, and with *Subject* as a blocking factor. A model was also run with the intensity change from the second vowel as the dependent variable but as the results are not meaningfully different, it is omitted here.

	Wald ChiSquare	*df*	Sig
(Intercept)	113.593	1	<0.001
Vowel_Environment	16.989	4	0.002
V#(j)V	6.384	1	0.012
Vowel_Environment*V#(j)V	8.088	4	0.088
Subject	98.275	46	<0.001

Table 3. GLiM with DV=*Intensity_change_from_V1*

Given that the intensity difference is statistically significantly greater in V#jV sequences than it is in V#V sequences, we see the intensity drop that is consistent with the idea that a phonological [j] would have a greater degree of constriction than an acoustic [j]. However, we see in the box plots in Figure 15 that there is a very large amount of overlap in the intensity change in V#V and V#jV forms, meaning that intensity drop is unlikely to be a reliable perceptual cue. Further, we see that Vowel Environment 5 (Short Classic [j]) shows a smaller intensity change in V#jV and a larger change when hiatus is present. While the reverse finding for Vowel Environment 5 may well be due to the very few modal V#jV tokens, we certainly at least see large-to-complete overlap in the ranges of the intensity changes with phonological versus acoustic [j]s across the Vowel Environments.

3.4. DURATION. Each V#(j)V sequence in Dataset 2 (as described in §2.4) had a measurement of duration taken. We would expect V#jV sequences to be longer than V#V sequences as they contain an additional segment, and, looking at the box plot in Figure 16 we see that sequences with a phonological glide are often longer than those without.

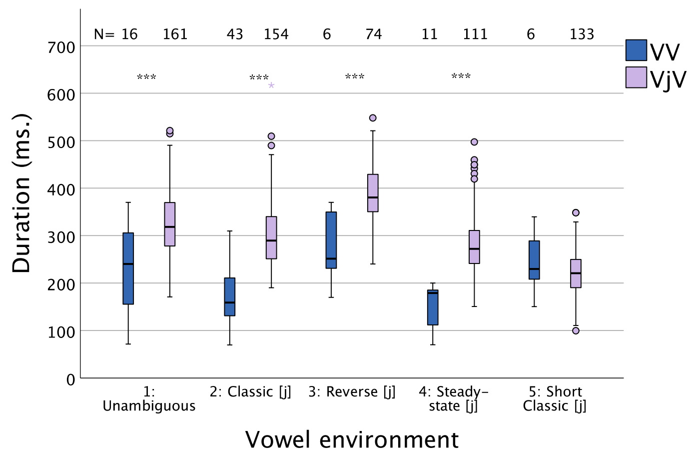

Figure 16. Duration of V#(j)V sequences in Dataset 2

In an attempt to increase the number of V#V tokens, the durations in Dataset 3, which included creaky voice tokens with trackable formants (see description in §2.4), were compared for modal voice and creaky voice. A GLiM of the V#V tokens (DV=*Duration*, IVs=*Voice*, *Vowel_Environment*, their interaction; with blocking by *Subject*) found that Vowel Environments 1 (unambiguous), 3 (Reverse [j]), and 5 (Short Classic [j]) did not significantly differ in duration between modal and creaky voice. Since Vowel Environments 2 (Classic [j]) and 4 (Steady-state [j]) did show a significant difference, they are excluded from the analysis. This subset of Dataset 3 (comprising Vowel Environments 1, 3, 5) contains 591 tokens. Figure 17 shows box plots of these durations, and Table 4 shows the statistical model results.

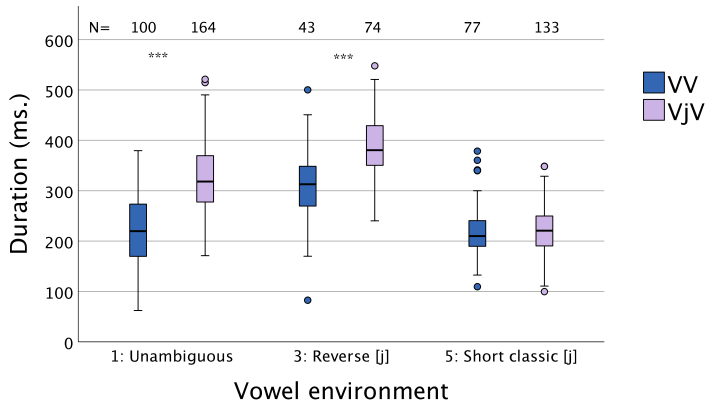

Figure 17. Duration of V#(j)V sequences in subset of Dataset 3

	Wald ChiSquare	*df*	Sig
(Intercept)	10806.288	1	<0.001
Vowel_Environment	347.955	2	<0.001
V#(j)V	109.524	1	<0.001
Vowel_Environment*V#(j)V	83.473	2	<0.001
Subject	111.554	46	<0.001

Table 4. GLiM with DV=*Duration*, V#(j)V sequences in subset of Dataset 3

Looking at the box plots in both Figure 16 and in Figure 17, we see that V#jV sequences are usually, but not always, reliably longer on average, and this is borne out by the statistical model. A pairwise comparison of V#jV to V#V sequences within each Vowel Environment finds those with a phonological glide are statistically significantly longer ($p < 0.001$) in Vowel Environments 1 (unambiguous) and 3 (Reverse [j]) but are not in Vowel Environment 5 (Short Classic [j]; $p = 0.898$). However, even looking at the Vowel Environments that show a robust difference, we again see a large overlap in durations of V#jV and V#V sequences, such that it would be hard for duration to be strong perceptual cue to the presence of a phonological glide.

3.5. FORMANT MOVEMENTS. F1 and F2 were extracted at 20%, 40%, 50%, 60%, and 80% through each V#(j)V sequence of Dataset 4 (as described in §2.4), which comprised all the modal voice token of Dataset 2 except for those which had non-trackable formants. The trajectory of each F1 and F2 was calculated as a summation of the Euclidean distance between each pair of adjacent measurement points. We might expect to see a Euclidean distance difference in Vowel Environment 1, where the lack of a [j] is not potentially confusable with the presence of a [j], however, based on the visual inspection of the trajectory summaries we saw that Vowel Environment 2 (Classic [j]) showed a difference between V#jV and V#V sequences as well. As can be seen in the box plots in Figure 18, V#jV involves more movement than V#V at least in Vowel Environments 1 and 2.

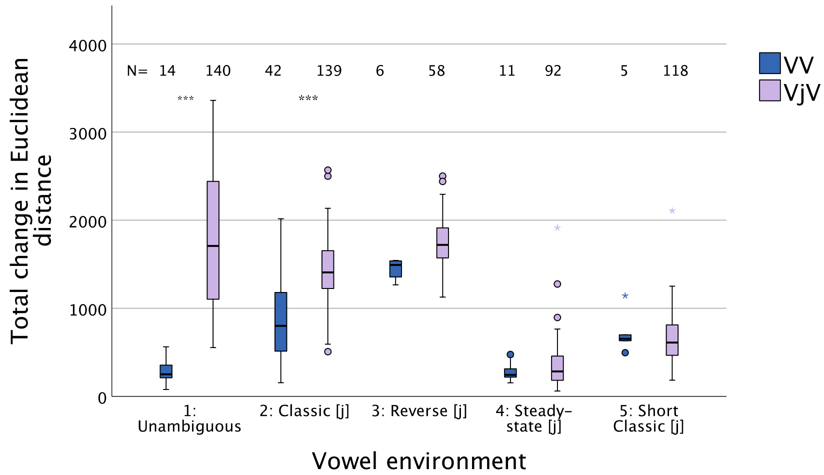

Figure 18. Summation of Euclidean distance between each F1, F2 measurement points in V#(j)V

The results of a Generalized Linear Model run in SPSS with dependent variable *Euclidean_Distance* and independent variables *Vowel_Environment*, *V#(j)V*, their interaction, and with *Subject* as a blocking factor are shown in Table 5.

	Wald ChiSquare	*df*	Sig
(Intercept)	752.977	1	<0.001
Vowel_Environment	179.866	4	<0.001
V#(j)V	44.163	1	<0.001
Vowel_Environment*V#(j)V	70.773	4	<0.001
Subject	88.710	46	<0.001

Table 5. GLiM with DV=*Euclidean_Distance*

Pairwise comparisons show that the difference between sequences with a phonological glide and those without are statistically significantly different in Vowel Environments 1 (unambiguous) and 2 (Classic [j]) ($p < 0.001$) but not in any of the others ($p \geq 0.276$). Even while the average Euclidean distance is clearly greater for V#jV sequences in Vowel Environment 2 (shown in Figure 18), its range overlaps almost completely with those of V#V sequences, again meaning that it would be hard for it to be used as a perceptual cue to the presence or absence of a phonological glide.

4. Conclusion. We have seen that the vocalic environment has an extreme effect on the pronunciation of [j]. In four of the examined Vowel Environments, a hiatus sequence is articulated much like the same sequence with a phonological glide. The comparisons of intensity change within the sequence, duration, and total Euclidean distance have supported that V#jV sequences are quantifiably different, on average, from V#V sequences. The greater intensity change is expected due to the consonantal nature of [j]. The longer duration is consistent with there being an additional phonological segment. Among the ambiguous Vowel Environments we only found greater Euclidean distance in Vowel Environment 2 (Classic [j]), which means speakers started the first vowel in the sequence from a higher, fronter position and ended the second vowel lower and backer when there was a phonological [j] in the sequence. Speakers could conceivably do this to cause the phonological [j] to have a more dramatic rise in F1 and fall in F2 than would occur in a V#V sequence. However, this was done only in one of the four ambiguous Vowel Environments and the differences were not extreme enough to allow a plausible categorization of any given Classic [j] V#(j)V sequence as one type or the other.

In all three measures, V#jV and V#V sequences exhibit such extreme overlap that they would appear to be ambiguous most of the time, at least based on that single measure. Further, we saw that Vowel Environment 5 (Short Classic [j]) did not exhibit any of the differences other ambiguous Vowel Environments did. Vowel Environment 5, then is potentially a case of full neutralization. The other three ambiguous cases appear to be cases of near neutralization, as typically the formant transitions that mimic a [j] are not measurably different from an actual, consonantal [j], even though, on average, they exhibit the expected differences in intensity change and duration.

The next stages of this project will pursue evidence from the perceptual side with English speakers, as well as production and perception evidence from other languages. Korean has minimal pairs such as [ki.ʌ] 'crawl' and [ki.jʌ] 'contribution', and Mandarin has two command particles [-ɑ] and [-jɑ], which are not in complementary distribution for at least some younger speakers. Future research will be able to speak to whether we find factors such as intensity change and duration being used to a more extreme degree in such cases.

A. Appendix: Stimuli. The following are the complete list of answer sentences. Each was read following the reading of the related question that introduced two possible endings to the answer sentence (see example (1)).

(2) Vowel environment 1: Unambiguous

a. A: The narrator will say that **Pa honors** his vows to his family.
b. A: Its the case that **Pa yawns** every evening.

c. They used the **pizza app** to order their dinner.
d. The childs toy **pizza yapped** to her friends toy elephant.

e. The syllabus specified an **extra essay** for the final.
f. The plan is for the director to use extra Johnson and **extra Yes** in the scene thats set outside the school.

g. The girl will see the **yellow oak** in the forest.
h. The girl will see a broken egg with a **yellow yolk** when she opens the styrofoam egg carton.

(3) Vowel environment 2: Classic [j]

a. The claim is that **she ought** to go to the party on the beach.
b. The story is that **she yachts** on the lake every week.

c. The announcement will say the event will **be all** day in the gym.
d. The narrator will say the leader claimed that it would **be yall** for the concert.

e. The narrator will say the boy didnt leave **me any** chocolate-covered raisins.
f. The voiceover will say that in Japan they gave **me yen** at the airport.

g. The children found the **ugly oak** in the courtyard.

h. She got the egg with an **ugly yolk** from the grocery market.

(4) Vowel environment 3: Reverse [j]

a. The narrator will say that **Ma eats** lunch at one.

b. The voiceover will say that **Ma yeets** the empty bottle into the recycling. [*yeet* slang for 'throw']

c. The locals will shout **hurrah ales** upon visiting the brewery.

d. The students will shout **hurrah Yale** at the baseball game.

(5) Vowel environment 4: Steady-state [j]

a. The story will be that **he eats** an orange every day.

b. The narrative will say **he yeets** the soda can down an alleyway.

c. She wanted to **see East** Berlin on foot.

d. The narrator will say the chef will **see yeast** rise by checking on it hourly.

e. His mother said **the east** road goes to Grandmas house.

f. They gave **the yeast** for the bread to the baker.

(6) Vowel environment 5: Short Classic [j]

a. It is evidenced that an **extra ear** can enhance auditory perception.

b. The advisor said the graduate program requires an **extra year** of research in the lab.

c. The farmers have to shuck **every ear** of corn in the fall.

d. Its the case that the class struggles **every year** on that quiz.

e. The traveller will see the **only inn** along the side of the road.

f. The girl will see the **only yin** yoga instructor in class.

References

Boersma, Paul & David Weenink. 2023. Praat: A doing phonetics by computer. Computer program. Version 6.3.10, retrieved 29 May 2023. http://www.fon.hum.uva. nl/praat/.

Davidson, Lisa & Daniel Erker. 2014. Hiatus resolution in American English: The case against glide insertion. *Language* 90(2). 482–514.

Lennes, Mietta. 2003. Praat script. Modified by Dan McCloy, December 2011; Retrieved 14 January 2016.

McCloy, Daniel R. 2016. *phonR: Tools for phoneticians and phonologists*. R package version 1.0-7. https://cran.r-project.org/web/packages/phonR/.

R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard & Jiahong Yuan. 2014. FAVE (forced alignment and vowel extraction) suite version 1.1.3. https://doi.org/10.5281/zenodo.9846.

RStudio Team. 2022. *RStudio: Integrated development environment for R*. RStudio, PBC. Boston, MA. http://www.rstudio.com/.