Hiatus resolution in American English: The case against glide insertion

Lisa Davidson, Daniel Erker

LANGUAGE

A JOURNAL OF THE LINGUISTIC SOCIETY OF AMERICA

Language, Volume 90, Number 2, June 2014, pp. 482-514 (Article)

Published by Linguistic Society of America
DOI: 10.1353/lan.2014.0028

HIATUS RESOLUTION IN AMERICAN ENGLISH: THE CASE AGAINST GLIDE INSERTION

LISA DAVIDSON
New York University

DANIEL ERKER
Boston University

It has generally been assumed that after nonlow vowels in English, hiatus is resolved by inserting a homorganic glide (e.g. *seeing* [sijɪŋ], Itô & Mester 2009). However, despite suspicions that inserted glides may be fundamentally different from lexical glides (e.g. Cruttenden 2008), a systematic phonetic investigation of the purported glide is lacking. We examine the nature of hiatus resolution by comparing three environments: (i) vowel-vowel sequences within words (VV: *kiosk*), (ii) vowel-vowel sequences across word boundaries (VBV: *see otters*), and (iii) vowel-glide-vowel sequences across word boundaries (VGV: *see yachts*). The first finding is that a glottal stop produced between the vowels accounts for nearly half of the responses for VBV phrases, whereas glottal stops are present in less than 5% of the VV and the VGV conditions. Second, an acoustic comparison of VV, VBV, and VGV phrases not produced with glottal stops shows significant differences between the vowel-glide-vowel and the vowel-vowel sequences on all measures, including duration, intensity, and formants. These results indicate that American English speakers tend to resolve hiatus at word boundaries with glottal stop insertion, whereas there is no hiatus resolution at all within words. A brief optimality-theoretic analysis sketches out the phonological differences between hiatus at word boundaries and word-medially.*

Keywords: hiatus resolution, American English, glottal stops, glides, acoustic analysis

1. INTRODUCTION. There is general agreement that for most vowel-vowel environments in English, hiatus is resolved with glide insertion. This assumption is so widespread that it has not merited discussion in the phonological literature per se; instead, the phenomenon of glide insertion is mostly mentioned in passing in studies that focus on the unusual phonological process of /ɹ/-insertion (e.g. Baković 1999, Broadbent 1991, Gick 1999, Harris 1994, Itô & Mester 2009, Krämer 2008, McCarthy 1993, McMahon 2000, Orgun 2001, Uffman 2007). The phenomenon of /ɹ/-insertion is limited to very particular environments, namely, following vowels that are not diphthongal, such as [ə, ɔ, ɑ], as in *law and order* [lɔɹənd] (Uffman 2007). In all other environments, there is ostensibly a glide insertion process to resolve hiatus. More specifically, it has been claimed that hiatus is resolved with insertion of /j/ when the first vowel is high and front, or has a high front offglide—for example, [i, eɪ, aɪ]—and with /w/ when the second vowel or offglide is high and back—for example, [u, oʊ, aʊ]. It is also generally assumed that glide insertion resolves hiatus in these vowel environments both across and within words (e.g. McCarthy 1993). Examples are shown in 1.

(1)	WITHIN WORDS	ACROSS WORD BOUNDARIES	
	seeing [sijiŋ]	see Ed [sijɛd]	(from McCarthy 1993)
	mosaic [moʊzeɪjɪk]	clay otters [kleɪjɑɾɚz]	
	coalition [koʊwəlɪʃən]	slow operas [sloʊwɑpɹəz]	
	fluid [fluwɪd]	new image [nuwɪmɪdʒ]	

Despite the prevalent assumption in the phonological literature of glide insertion as hiatus resolution, a few authors have noted, at least impressionistically, that glides inserted in a hiatus-resolution context may not be acoustically identical to lexical /j/ and

---

* We would like to thank Maria Gouskova and Gillian Gallagher for their extensive comments on this work. We would also like to thank members of the NYU PEP Lab, the colloquium audience at UCSC, the audience at the 2013 LSA annual meeting in Boston, and the Spring 2012 Acoustic Phonetics class at NYU for their feedback.

/w/. One early account by Stene (1954:19) notes that hiatus resolution may not occur when vowels are preceded by high front or back vowels or offglides because 'they have as a transitional element a near-consonantal [j-] or [w-] glide, which is a sufficient closure for the onset of the next vowel'. In later work, Cruttenden (2008) observes that 'linking [^{j,w}]' are not the same as phonemic /j, w/, and that there are minimal pairs that illustrate the difference between inserted and lexical glides: *my ears* vs. *my years* and *two-eyed* vs. *too wide*. Heselwood (2006:80) goes one step further and refers to [^{j,w}] as 'low-level articulatory transitional phenomena' (see also similar descriptions in Britain & Fox 2008, Newton & Wells 2002). Moreover, a perceptual study by Hunt (2009) demonstrates that listeners can distinguish between synthesized /i#ji ~ i#i/ and /u#wu ~ u#u/ sequences. A similar claim has been made for Dutch hiatus contexts across word boundaries (van Heuven & Hoos 1991), where acoustic evidence provides partial support for distinguishing lexical glides from glides said to resolve hiatus (i.e. there were substantial differences in duration, but not in formant measures). What these descriptions and studies have in common is that they do not treat glide insertion as a phonological process to break up sequential vowels, but rather as epiphenomenal to an articulatory transition between two vowels. Yet, at least for English, the descriptions of the glide as a transitional element are speculative and do not rely on any phonetic evidence that bears on the claims.

Another line of research suggests that there are at least some environments in English in which a glottal stop may surface between two sequential vowels, at least at the beginning of a vowel-initial word. Studies examining the distribution of glottal stops in English have found that they generally occur before vowel-initial words, and that their presence correlates with the strength of the prosodic boundary (Dilley et al. 1996, Garellek 2012a,b, Pierrehumbert 1995, Pierrehumbert & Frisch 1997, Pierrehumbert & Talkin 1992, Redi & Shattuck-Hufnagel 2001). For example, using a corpus of radio announcer speech, Dilley and colleagues (1996) found that the presence of a glottal stop or glottalization on vowel-initial words was relatively high for an initial vowel that was also utterance-initial, regardless of which phoneme preceded the pause (40–50%). In phrase-medial position, however, there was a substantial difference between glottalization when the preceding segment was a vowel (approximately 30% glottal stop presence) and when it was a consonant (≤ 20% glottal stop presence). Only one study, by Mompeán and Gómez (2011), specifically examined the use of glottal stopping as a hiatus-resolution strategy in British English, though their environments were limited to those that were candidates for /ɹ/-insertion since that was the focus of their study. Mompeán and Gómez found that 31.1% of possible /ɹ/-insertion environments across word boundaries were instead produced with a glottal stop. Thus, there is some evidence that at least across word boundaries and in British English, a glottal stop is used to resolve hiatus. However, no similar study has been carried out for word-internal vowel-vowel sequences.

Crosslinguistic research suggests that another factor that may condition glottal stop insertion is the stress pattern of the following syllable. The literature on glottal stop production in German, for example, has shown that it is more likely to occur before stressed, vowel-initial syllables, though it can also be found before unstressed syllables (Alber 2001, Kohler 1994, Wiese 1996 for German, and Jongenburger & van Heuven 1991 for experimental results on a similar process in Dutch). Mompeán and Gómez (2011) present similar results across word boundaries in a subset of their data. These results suggest that since a stressed syllable is a stronger prosodic unit than an unstressed syllable, hiatus resolution may be more likely to occur in that environment; the current study considers this possibility in detail.

Although the phenomenon of hiatus resolution in English has featured prominently in the phonological literature, it has received almost no attention in the phonetic literature. Yet, it is clear from the studies reviewed above that the phonetic properties of whether and/or how hiatus is resolved are unknown. Thus, the goal of the current study is to present an acoustic investigation of vowel-vowel sequences to determine whether glide insertion actually is a process of hiatus resolution in American English. In this study, we limit our investigation to /j/- and /w/-insertion environments, since our main question focuses on the status of glide insertion and not on the acoustic properties of the rhotic in dialects that have /ɹ/-insertion. To preview the results, our findings suggest that the extent of hiatus resolution in English has been simultaneously overestimated and mischaracterized. In particular, we find that glide insertion does not occur following a vowel with a homorganic offglide, and glottalization plays a significant role in hiatus resolution, especially across word boundaries and before stressed syllables.

This study aims to investigate two main questions about hiatus in English. First, is there acoustic evidence to support the claim that English resolves hiatus? Second, if there is evidence for hiatus resolution, is it the same process across different phonemic and prosodic contexts? To address these questions, we focus on three types of utterances: word-INTERNAL vowel-vowel sequences (VV: e.g. *kiosk*), word-BOUNDARY vowel-vowel sequences (VBV: e.g. *see otters*; ‘B’ refers to ‘boundary’), and vowel#glide-vowel sequences (VGV: e.g. *see yachts*). Since there are several potential ways for hiatus resolution to occur in English, in 2 we hypothesize four possible scenarios to account for how English speakers may produce these three types of utterances. These predictions are stated in general terms; the details of the acoustic analysis are provided in §2.1.

(2) a. English resolves all hiatus with glide insertion. If this scenario is true, the acoustic properties corresponding to the glide should be the same for VV, VBV, and VGV sequences.
  b. English allows hiatus across the board; apparent ‘resolution’ is attributable to the offglide of the vowel. If this scenario is true, the acoustic properties of VV and VBV should be the same, but they should be different from those of VGV.
  c. VBV is resolved with a glottal stop, but glide insertion is the resolution strategy for VV. In this case, the acoustic properties of VV and VGV would be the same, but different from VBV.
  d. VBV is resolved with a glottal stop, but hiatus is allowed in VV sequences. In this outcome, there should be substantial acoustic differences between VV and VGV sequences.

In the following sections, we detail both the categorical and the continuous measures that were used to analyze the VV, VBV, and VGV sequences.

2. Experiment.

2.1. Methods.

Participants. Speakers included fourteen monolingual participants (seven men, seven women) recruited through New York University undergraduate classes, ranging in age from eighteen to twenty-five. The participants reported growing up in the following locations: Georgia, Philadelphia, Pittsburgh, New Jersey, Minnesota, New Mexico, New York City, upstate New York, Austin, TX, Massachusetts, Chicago, and

Michigan.[1] Thirteen participants reported no history of speech or hearing disorders and one participant reported speech therapy as a small child for the misarticulation of /s/. Since this was unrelated to the question being studied, this participant's data was retained. The listeners were compensated for their time. No participant reported learning another language before age eight.

Materials. The target stimuli included twenty-four words and phrases in each of three categories: a vowel-vowel sequence within a word (VV; e.g. *kiosk* /iɑ/, *stoic* /oɪ/, *duo* /uo/), a vowel-vowel sequence across a word boundary (VBV; e.g. *see otters* /i#ɑ/, *know itchy* /o#ɪ/, *two oboes* /u#o/), and a vowel-glide-vowel sequence across a word boundary (VGV; e.g. *see yachts* /i#jɑ/, *go witness* /o#wɪ/, *Sue woke* /u#wo/). For each of the VV, VBV, and VGV categories, there were twelve vowel combinations, six consistent with the glide or offglide /j/ (/i-ɑ,ə/, /e-ɑ,ə/, /i-ɪ/, /e-ɪ/, /i-o/, /e-o/) and six consistent with /w/ (/o-ɑ,ə/, /u-ɑ,ə/, /o-ɪ/, /u-ɪ/, /o-o/, /u-o/). A list of the stimuli is given in Appendix A. The words or phrases were designed such that each vowel sequence was flanked by an obstruent or nasal on either side in almost all of the stimuli in order to ensure that segmentation would be relatively straightforward for the acoustic analyses. Since it was relatively difficult to find pairs of words with the vowel sequences in question, frequency of the words was not controlled during the selection of the stimuli; however, frequency was factored into the analyses of both categorical and continuous variables in §3. Likewise, the stress patterns of the stimuli, especially of the second vowel in the sequence, was initially not taken into consideration in the design of the study and was not controlled. However, since it became clear after coding the data that stress was a relevant factor, stress was included in the statistical analyses (described in detail below).

These words and phrases were incorporated into three reading passages. Care was taken to ensure that the words and phrases were not utterance- or phrase-initial or -final (with one exception, *indigo ink*). Otherwise, no other particular restrictions on the stimuli were imposed since the goal of this study was to get a general picture of hiatus-resolution strategies within a fluent context. The reading passages are given in Appendix B.

Procedure. Recordings took place in a sound-attenuated booth at New York University. Participants had been sent copies of the reading passages by email and were asked to familiarize themselves with the passages before coming to the recording session. Participants were asked to read each of the reading passages twice, once at a slow rate and once at a fast rate. These rates were not controlled by an external clock, but subsequent analysis confirmed that the passages were read significantly faster when participants were told to read at a faster rate (see §3.2).

Recordings were made with a Zoom H4n digital recorder at a sampling rate of 44 kHz and a head-mounted Shure WH30 microphone. Since the head-mounted microphone stays a fixed distance from the mouth during the recording, intensity analyses can be carried out.

2.2. Data analysis. Using Praat 5.2 (Boersma & Weenink 2011), the vowel-vowel or vowel-glide-vowel interval of each stimulus item was segmented in a textgrid. For intervals after stops and fricatives, segmentation began at the onset of F2. The end of

---

[1] Although our participants are from varied geographic regions, examination of the individual speakers shows that they all behave similarly on both the categorical responses and continuous measurements. It does not seem that dialectal variation is a factor in conditioning the responses in this study.

the interval was marked at the offset of F2 preceding stops and fricatives. Intervals before and after nasals began at the sharp change in intensity between the nasal formants and oral vowel formants.

Each stimulus item was categorically coded for one of the following properties. Examples of these coding types and of the segmentation criteria are given in Appendix C. There were also some responses that were coded as 'no response', where speakers did not produce the intended item or produced it with a pronounced disfluency, including a perceptibly unnatural pause between the two words in the VBV or VGV conditions.

- MODAL: modal voice throughout the vocalic interval
- GLOBAL CREAK: creak that lasted the entire duration of the interval
- V1 CREAK: creak only on the first vowel, or creak that lasted no more than the first half of the interval
- V2 CREAK: creak only on the second vowel, or creak that occurred only on the second half of the interval
- GLOTTAL STOP: a glottal stop between modal vowels. To be classified as a glottal stop, a period of silence had to be flanked on either side by fully modal vowels.
- GLOTTALIZATION: glottalization between modal vowels. To qualify as glottalization, the utterance had to include a period of irregular phonation (but no silence) flanked by fully modal vowels.

The distinction we are making between 'creak' and 'glottalization' primarily pertains to differences in where these properties are located, not to potential acoustic differences. As we argue below, speakers seem to be implementing the same acoustic characteristics for two different purposes, which we keep separate by using two different terms. The identification of creak/glottalization followed the descriptions in Dilley et al. 1996 and Redi & Shattuck-Hufnagel 2001. The main acoustic correlate of creak/glottalization was substantial irregularity in the duration and amplitude of glottal pulses from period to period (see examples in Appendix C) (Redi & Shattuck-Hufnagel 2001: 414). Cases were classified as glottal stops only if there was a discernible period of silence in the middle of vowel-vowel or vowel-glide sequences. Instances of creakiness or glottalization were typically very easy to identify, since the irregular phonation was distinct both in the shape of the glottal pulses and in intensity. The very few cases that were unclear were classified as modal, since this is the more conservative categorization. For the classification of V1 or V2 creak, approximately half of the whole vocalic interval had to show creak. Thus, cases that demonstrated a few milliseconds of creak immediately following the preceding consonant or before the following consonant were not considered to show V1 or V2 creak. Such tokens were classified as modal unless they also contained a glottal stop or glottalization (otherwise flanked by modal phonation), in which case they were classified as glottal stop or glottalization. Nevertheless, such cases were extremely rare.

The reason that so many different types of categorical patterns were coded was because we were interested in potential linguistic and nonlinguistic uses of creak and glottal stopping, and whether there is any evidence that these different categories could signal different phenomena. Recent literature suggests that some speakers, especially young women, tend to use creak as a regular feature of particular registers and not for explicitly phonemic purposes (Wolk et al. 2012, Yuasa 2010). We hypothesized that there might be differences in the distribution of the categorical variables, such that global creak or V2 creak could be more indicative of creak as a voice quality rather than a phonological element used in hiatus resolution. Impressionistically, both global creak and V2 creak occasionally persisted into neighboring segments (when such segments were not obstruents), but, by definition, this was not the case for glottal stops and glottalization. Rather, glottal stops or periods of glottalization between two modal vowels have previously been shown to be indicative of hiatus resolution (e.g. Mompeán & Gómez 2011, Pierrehumbert 1995, Pierrehumbert & Frisch 1997). As is seen below, results suggest that it is crucial to distinguish between 'creak' and 'glottalization', such that the former is used to refer to nonsegmental uses of creakiness, whereas the latter refers to implementations that relate to hiatus resolution.

For just those vowel-vowel and vowel-glide sequences that were labeled as 'modal', several acoustic measurements were also taken. These measurements were intended to examine whether the acoustic properties of the vowel-vowel and vowel-boundary-vowel sequences (VV and VBV) were indistinguishable from those of the vowel-glide-vowel sequences (VGV), which would indicate that speakers do in fact insert a glide as another method of hiatus resolution.

- DURATION MEASUREMENTS: the duration of the entire vocalic or vowel-glide-vowel interval was measured. This is a very rough metric, since we did not control for speech rate or for the specific environment of the target words (except that they were neither phrase-initial nor phrase-final). However, we at least expect that the durations of VV and VBV should be significantly shorter than for VGV if there is no glide insertion. We note that the duration measure provides only converging evidence, and that more weight should be put on the intensity and formant measures.
- INTENSITY MEASURES: Within the vowel-vowel or vowel-glide-vowel sequence, three intensity measures were taken: minimum RMS intensity occurring between 25% and 75% of the duration of the interval (intensity min), maximum RMS intensity in the period from the onset of the first vowel to the minimum intensity (V1 intensity max), and maximum RMS intensity in the period between the minimum intensity and the offset of the second vowel (V2 intensity max). We assumed that the intensity minimum in the VGV sequences corresponded to the maximal achievement of the glide or diphthongal offglide, but we restricted the measure to the middle portion of the vowel to ensure that we did not get spurious measures corresponding to the ramping up of intensity for V1 or a decrease for V2. Two difference scores were then calculated: V1 intensity max − intensity min, and V2 intensity max − intensity min. We hypothesized that both of these differences should be significantly greater for VGV than for either VBV or VV if there is no glide insertion for hiatus resolution.
- FORMANT MEASURES: Two different formant measurements were taken depending on whether the glide (or potential inserted glide) was /j/ or /w/. Both analyses were conducted using linear predictive coding in Praat, with a window length of 25 ms and a maximum formant specification of 5500 for the female speakers and 5000 for the males. For /j/, we measured the F2 value (following Aguilar 1999) at the time of the intensity minimum, and for /w/, we measured the F1 value at the intensity minimum. Though we also intended to measure F2 for /w/, there were a number of cases in the VGV stimuli where the amplitude of F2 was attenuated substantially and could not be properly measured by Praat (i.e. no values were reported in the output of the Praat script). Since F1 is also lowered in labial sounds like /w/ (e.g. Espy-Wilson 1992), we used it instead (see also Hunt 2009 on F1 lowering in glides). We hypothesized that F2 should be higher in VGV sequences than in either VBV or VV if there is no glide insertion, and that F1 should be lower.

3. RESULTS.

3.1. CATEGORICAL VARIABLES. The analysis of each of the categorical variables defined in §2.1 (modal, global creak, V2 creak, glottal stop, glottalization) was carried out with a mixed-effects binomial logistic regression implemented in R (R Development Core Team 2012) with the lme4 package (Bates & Sarkar 2008). We use mixed-effects models (MEMs) for both the categorical variables and the continuous variables in §3.2 for two reasons. First, standard analyses of variance (ANOVAs) are not appropriate for binomial data, and mixed-effects binomial regressions provide more accurate analyses of the data than arcsine-transforming the data (see arguments in Jaeger 2008, Quené & van den Bergh 2008), which was the commonly recommended fix for making binomial data better conform to the assumptions of ANOVAs. Second, it is commonly accepted that MEMs address the problem of random effects in linguistic data (Baayen 2008, Baayen et al. 2008). That is, statistical models are more accurate if subjects and items are both treated as random factors, since models with random factors better generalize to stimuli and participants beyond those in the study itself. Because in any given analysis ANOVAs can treat only subjects or items as a random factor, but not both, MEMs are more powerful and accurate statistical analyses.

Each analysis was carried out separately; that is, we investigated the effect of the independent variables on each of the categorical dependent variables in turn. The analysis was carried out by comparing each of the categorical dependent variables to the aggregation of all other responses. Although our data might be alternatively accounted for with a multinomial model that includes each of the possible response types as part of the dependent variable, typical implementations of multinomial models do not allow for mixed-effects modeling. Following much recent precedent in the literature, we maintain that the advantages of mixed-effects models outweigh those offered by a multinomial model, and so we examined each response type separately. However, we include a multinomial model in Appendix D to verify that the same results are also obtained using that method. For each of our binomial models, we first attempted a 'maximal' random-effects structure for the by-participant effects including random slopes. But since most of our models did not converge with the full random-effects structure for both by-participant and by-item effects, we simplified our models in ways that allow for sensible interpretation of the data.[2]

In each of these binomial analyses, responses that were coded for the categorical variable being analyzed are the dependent variable. The independent variables are gender (male, female), rate with which the passage was read (slow, fast), the type of vocalic sequence (VGV (glide), VBV (word boundary), VV (word-internal)), and frequency of the word in the case of VGV or the second word for VV and VBV items. Log frequency was calculated using the SUBTLEX$_{\mathrm{WF}}$-US database implemented on the English Lexicon Project website (Balota et al. 2007, Brysbaert & New 2009). For the VV and VBV items, the second word was chosen for the frequency count because the first word was often a pronoun, numeral, or intensifier (e.g. *he*, *too*, *two*), whereas the second word was usually a content noun or verb. In addition, it was our impression listening to the

---

[2] For the categorical variables, the most complex model that converged for the sequences coded as 'modal' was of the form: (modal ~ gender + rate + hiatustype + logfreq + (1 + gender + rate + logfreq|words) + (1 + rate + hiatustype + logfreq|subject), family=binomial('logit')). Because the model in which all of the effects were fully crossed did not converge, we decided to use the model with main effects only rather than use the fully crossed model with only random intercepts (for further information, see n. 3). For the remaining dependent variables, which had fewer observations, the by-items (word) random effect had to be simplified to the random intercept only in order for the model to converge. The by-subjects effect remained maximal.

files that speakers always correctly produced the first word in the sequence, but occasionally stumbled when speaking the second word. Thus, we were interested in whether the frequency of the second word affected whether speakers produced it with modal voice or with some other quality.

Stress of the second vowel is not included in the initial statistical models because of an imbalance in the data set. For the vowel-glide-vowel (VGV) category, none of the stimuli were unstressed on the second vowel. For the word-internal (VV) stimuli, 25% of the words (six stimuli) had stress on the second syllable, whereas 75% of the stimuli (eighteen items) in the word-boundary category (VBV) had stress on the second syllable. Since the VGV category does not have any stimuli in which the second syllable of the sequence is unstressed, stress cannot be included as a factor in the full models reported below. Thus, stress is treated separately for a subset of the data after the initial analyses are reported.

The proportions for each categorical code are shown for gender and speech rate in the graph in Figure 1. There were 1,987 total tokens, and the overall proportions for each of the categorical variables are shown in Table 1. Because V1 creak items constitute such a small proportion of the data, they are not further considered in the statistical analysis below. In addition to the detailed graph in Fig. 1, the responses are also broken down by sequence type alone in Table 1.

	modal	creak	glot	glot stop	creakV2	creakV1	no resp
VBV	34%	5%	28%	17%	12%	2%	2%
VGV	76%	2%	4%	1%	12%	1%	5%
VV	74%	5%	2%	0%	13%	2%	4%
overall	**61.3%**	**3.7%**	**11.2%**	**6%**	**12.2%**	**1.8%**	**3.8%**

Table 1. Percentages of responses for each sequence type. (For key to abbreviations, see Fig. 1.)

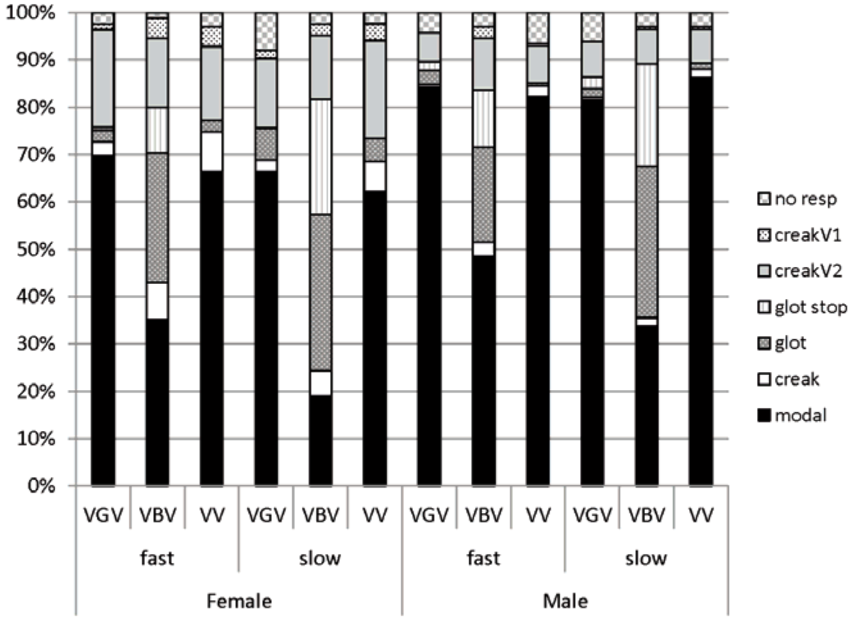

Figure 1. Percentage of each categorical response type divided by sequence type, rate, and gender. Abbreviations: no resp: no response, creakV1: creak on V1 only, creakV2: creak on V2 only, glot stop: glottal stop, glot: glottalization between modal vowels, creak: global creak on whole sequence, modal: modal production of vocalic sequence.

MODAL RESPONSES. The results of the binomial regression for the responses coded as modal are given in Table 2. For this and all other analyses of categorical variables, the reference variables are VGV, fast rate, and female. For binomial variables, the 'estimate' column in the statistical output (also referred to as β) is the likelihood of change in log odds as compared to the reference variables. Negative estimates indicate that the relevant independent variable decreases the likelihood of the dependent variable, whereas variables with positive estimates increase the likelihood. These results show that males are significantly more likely to have modal responses than females, slow tokens are less likely to be modal than fast tokens, and VBV is less likely to be modal than VGV, but there is no significant difference between VGV and VV. Post-hoc Tukey tests using the multcomp package in R (Bretz et al. 2010) show that VV is significantly more likely to be modal than VBV (β = 3.40, $z = 5.99$, $p < 0.001$). There is no significant effect of frequency.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	1.441	0.538	2.679	0.007*
rate: slow	−0.565	0.167	−3.380	0.001*
seq: VBV	−3.446	0.523	−6.590	0.000*
seq: VV	−0.014	0.453	−0.030	0.976
frequency	0.157	0.207	0.758	0.449

TABLE 2. Results for modal responses. In this and following tables, the baseline values for the independent variables are female for gender, fast for rate, and VGV for sequence type.

GLOBAL CREAK. The results of the regression for the global creak responses are given in Table 3. These findings show that males are significantly less likely to have creaky responses than females, and that VBV is significantly more likely to be creaky than VGV (and VV marginally so). There is no significant difference between VV and VBV (β = −0.64, $z = -0.726$, $p = 0.47$). Neither rate nor frequency are significant.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	−1.328	0.631	−2.105	0.035*
rate: slow	−0.169	0.307	−0.550	0.582
seq: VBV	2.083	0.974	2.138	0.033*
seq: VV	1.359	0.749	1.813	0.070(*)
frequency	−0.562	0.343	−1.640	0.101

TABLE 3. Results for global creak ('creak') responses.

V2 CREAK. The results of the regression for creak on the second vowel are given in Table 4. The only finding, that males are less likely to have creak on V2 than females, is approaching significance.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	−0.595	0.317	−1.874	0.061(*)
rate: slow	−0.077	0.156	−0.492	0.623
seq: VBV	0.255	0.570	0.448	0.654
seq: VV	0.562	0.522	1.077	0.281
frequency	0.170	0.222	0.767	0.443

TABLE 4. Results for creak on V2 ('creakV2') responses.

GLOTTAL STOPS. The findings for responses with glottal stops are presented in Table 5. These results show that glottal stops are significantly more likely in VBV sequences than they are in VGV, and that they are more likely at slower rates of speech. Because glottal stops are never produced in VV, the appropriate estimates cannot be obtained. This suggests that VBV, which has glottal stops in 17% of the utterances, is different from VV (and is significantly different from VGV, as shown in Table 5). VGV has glottal stops in 1% of the utterances.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	−0.381	0.575	−0.662	0.508
rate: slow	1.107	0.326	3.396	0.001*
seq: VBV	4.378	0.715	6.120	0.001*
frequency	−0.399	0.284	−1.403	0.160

TABLE 5. Results for glottal stop ('glot stop') responses.

GLOTTALIZATION. The findings for responses with a period of glottalization between modal vowels are given in Table 6. These results show that glottalization is significantly more likely to occur in VBV sequences than in VGV, but significantly less likely in VV. Glottalization is also significantly less likely to occur in VV as compared to VBV ($\beta = -3.97$, $z = -6.32$, $p = 0.001$). Glottalization is also more likely to occur at the slow rate than the faster rate.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	−0.448	0.307	−1.460	0.144
rate: slow	0.622	0.212	2.935	0.003*
seq: VBV	2.881	0.471	6.122	0.000*
seq: VV	−1.096	0.557	−1.966	0.049*
frequency	−0.240	0.202	−1.188	0.235

TABLE 6. Results for glottalization ('glot') responses.

COMPARISON OF CREAK AND GLOTTALIZATION. In order to further examine whether there are significant differences between the creak variables and the glottalization/glottal stop variables, we also carried out a binomial analysis on a subset of the data containing only these productions. The creak variables (V2 creak and global creak) were collapsed and were compared to the glottalization variables (glottalization and glottal stop). If there are significant differences in the independent variables affecting these two types of responses, it would further justify the conclusion that nonmodal voice quality is being used in two different ways in this data.

For this analysis, we include gender, speech rate, and sequence type as independent variables. Since this data set converges with fully crossed factors, we use the more complex model here, including random intercepts for participants and items. Main effects and significant interactions are shown in Table 7; the two- and three-way interactions that were not significant are not included in the table.

	ESTIMATE	STD. ERROR	*z*-VALUE	*Pr*(>|*z*|)
gender: male	1.734	0.976	1.776	0.076(*)
rate: slow	1.535	0.781	1.966	0.049*
seq: VBV	2.838	0.943	3.009	0.003*
seq: VV	−1.292	1.164	−1.110	0.267
gender: male * seq: VV	−3.638	1.839	−1.973	0.048*

TABLE 7. Results for a comparison of glottalization and creak variables.

Though the results in Table 7 are complex, they clearly corroborate the finding from above that creakiness and glottalization are differently conditioned phenomena. Specifically, they show that the likelihood that males produce more glottalization responses overall as compared to creak responses is marginally significant, but the significant interaction indicates that creaky responses for males for the VV sequence are even less likely. There are significantly more glottalization responses for the slow rate and for the sequence VBV as compared to VGV. VV is not significantly different from VGV, but there are significantly fewer glottalization responses as compared to creak responses for VV as compared to VBV ($\beta = -4.11$, $z = -4.18$, $p < 0.001$).

EFFECT OF STRESS. For analyses examining stress, the relevant comparison is between VBV and VV stimuli, since the VGV stimuli did not contain any items with stress on V2. Here we ask whether the stress of the second vowel in the sequence conditions the type of categorical response that is produced. For each of the previous analyses in §3.1, the same models as above were run for the subset of data containing only VBV and VV stimuli, except that they now also included a factor for stress on V2 and an interaction between sequence type (VBV or VV) and stress (stressed V2 or unstressed V2). For VV, there were 168 items with stress on V2 and 504 with no stress on V2. For VBV, there were 492 items with stress on V2 and 168 items with no stress on V2. In Table 8 we report results only for these new factors because the results for the other variables mirror what was reported in the previous sections.

		ESTIMATE	STD. ERROR	*z*-VALUE	$Pr(>\lvert z\rvert)$
MODAL	V2 stress	−2.61930	0.598	−4.378	0.0001*
	seq: VV * V2str	2.60670	0.794	3.282	0.001*
GLOBAL CREAK	V2 stress	−1.51360	1.060	−1.427	0.153
	seq: VV * V2str	0.00682	1.548	0.004	0.996
V2 CREAK	V2 stress	0.70370	0.703	1.001	0.317
	seq: VV * V2str	−1.22310	0.933	−1.311	0.189
GLOTTALIZATION	V2 stress	1.59147	0.549	2.901	0.004*
	seq: VV * V2str	−0.04220	0.957	−0.044	0.965
GLOTTAL STOP	V2 stress	1.70960	0.517	3.308	0.001*
	seq: VV * V2str	−1.57260	2543.760	−0.001	0.999

TABLE 8. Results for categorical responses including stress as a predictor.

The results of the analyses including stress demonstrate that there are significantly fewer modal responses when stress is on the second vowel of the sequence. The interaction for the modal responses indicates that there are fewer modal responses when V2 is stressed in VBV sequences (59.5% modal productions when V2 is unstressed versus 25.4% when it is stressed) as compared to the pattern for VV (74% modal production for unstressed V2 versus 75% for stressed V2). The significant results for glottalization and glottal stops indicate that there are significantly more of these responses when stress is on V2. The overall pattern for VV and VBV—where stress has the main effect—is shown in Figure 2. It is clear from this figure that glottalization and glottal stops are much more likely to occur before stressed vowels at word boundaries than word-internally.

SUMMARY: CATEGORICAL VARIABLES. The results for the categorical variables converge on the following findings. First, overall, 61.3% of the responses are modal productions, whereas 34.9% of the responses contain some kind of glottalization or creak. We chose not to collapse all of the creaky and glottalized responses into a single category because we believe that they constitute separate phenomena, and that speakers employ them for different purposes. Specifically, the results indicate that global creak and V2 creak are the only classes of response that are either significant or marginally significant

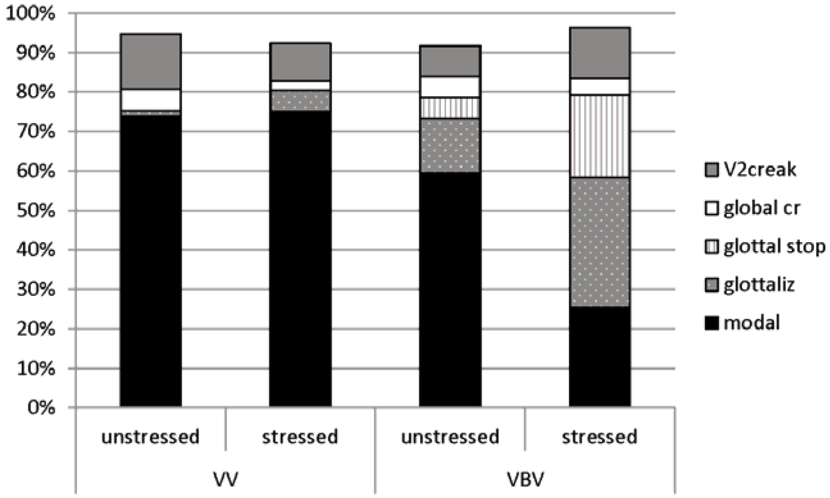

FIGURE 2. Percentage of responses in VV and VBV sequences for stressed and unstressed vowels in V2 position. The columns do not add up to 100% because responses like 'V1 creak' and 'no response' are not included.

for the gender variable, whereas there is no significant difference between males and females for glottal stops or glottalization. The comparison of creak variables to glottalization variables is generally consistent with this interpretation, since it shows that males have fewer creaky responses than glottalized responses as compared to females. Although we have not quantified total amounts of creakiness in the rest of the sound files, this finding is compatible with previous findings reporting that females produce more creakiness ('vocal fry') in general than males do (Wolk et al. 2012, Yuasa 2010). That is, global creak and V2 creak do not seem to be signaling a linguistic use within a very constrained location, whereas glottal stops and glottalization are limited to only a short period between two modally produced vowels. Likewise, rate is a significant predictor only for glottal stops and glottalization—which are more likely at the slower rate—but not for global creak or V2 creak. This may be a result of an articulatory cost associated with inserting a glottal articulation between modal vowels (Borroff 2005), which may be easier to produce at a slower rate. Global creak and V2 creak, however, are produced for a longer duration (and were impressionistically produced beyond the boundaries of the vowel-(glide-)vowel sequences that were coded for this study).

The results for the sequence type show that for almost every response type, VBV is significantly different from VGV. There are significantly fewer modal responses for VBV, and significantly more global creak, glottal stop, and glottalization responses for VBV than for VGV. Only V2 creak is not significantly different. The pattern for VV as compared to VGV is more varied; there is no significant difference in modal, V2 creak, or glottal stop responses, marginally more global creak responses, and significantly fewer glottalization responses. Together, these findings indicate that glottalization and glottal stop responses are a marker of hiatus resolution between words. The results from the analysis of stress additionally demonstrate that the environment for hiatus resolution can be further localized to preceding stressed vowels, indicating that there is a pattern of using glottalization to mark boundaries before stressed syllables. This is consistent with findings in other studies that examined a subset of the conditions examined in this study (Dilley et al. 1996, Garellek 2012a,b, Mompeán & Gómez 2011, Pierrehumbert 1995, Redi & Shattuck-Hufnagel 2001). Glottalization and glottal stops are almost never used to resolve the hiatus in VV sequences (from Table 1, 2% of the cases), whereas they are frequent in VBV sequences (45% of the cases). This is confirmed by the proportions for the modal responses: 34% of VBV sequences had a modal response, as compared to 76% for VGV and 74% for VV.

Based on these results and previous characterizations of the phonetic implementation of glottal stops (Esling & Harris 2005, Garellek 2012a,b, Ladefoged & Maddieson 1996, Ogden 2001), we argue that glottalization and glottal stops are poles on a continuum of phonetic implementation for the phonological category of glottal stop. Most importantly, both glottalization and glottal stops are realized as an interval of laryngealized phonation between two periods of modal phonation, indicating that it is precisely controlled and corresponds to a linguistic (presumably phonemic) use of glottalization. Thus, for the rest of the discussion, we treat glottalization and glottal stops as implementations of the same phonological category.

Finally, the results for the categorical variables reveal that there is no significant effect of frequency for any of the response types. This indicates that glottalization and glottal stops as hiatus-resolution strategies for VBV tokens are independent of frequency in this data set,[3] just as creaky responses, which were more strongly implemented in female speakers, do not seem to be affected by the frequency of the utterances.

In the next section, we turn to an acoustic analysis of the modal responses. Although glottalization and glottal stops do not seem to be a strategy for hiatus resolution for VV sequences, the analysis of the categorical variables does not rule out glide insertion as a method for hiatus resolution for the modal responses for VV or VBV.

3.2. Continuous variables. The analysis of the continuous variables (duration, intensity, formant values) was carried out with a linear mixed-effects regression implemented in R with the lme4 package. We followed the same convention of building maximal models and simplifying them when they do not converge that we did for the categorical variables. For the continuous variables, fully crossed models did converge, so interactions are reported in this section.

In the statistical tables for the continuous variables, the column labeled 'estimate' shows the predicted amount of change in the dependent variable for each independent variable with respect to the reference values in the model. A positive value means the independent variable would lead to an increase in the dependent variable (e.g. the slow condition leads to an increase in duration in milliseconds), whereas a negative value indicates a decrease in the amount of the dependent variable. Unlike the analysis of the categorical variables using a binomial regression, no *p*-values are returned for the analysis of continuous data. While there are techniques for determining *p*-values that have been suggested for linguistic data, such as Markov chain Monte Carlo sampling (Baayen 2008), this technique has not been implemented in lme4 for models that contain random correlation parameters. We instead follow the practice of using the *t*-statistic (the estimate divided by the standard error) as a diagnostic of whether the independent factors are a significant contributor to the model (Gelman & Hill 2006); if it is above 2 or below −2, then we consider the factor significant.

---

[3] Although the maximal model was simplified as explained in n. 2 in order to get sensible models that converged, we also carried out binomial regression models with fully crossed independent variables and random intercepts only, with the specific goal of examining whether there were interactions between frequency and the other independent variables. The results of those models indicated that not only was frequency never a significant main predictor, but it also never significantly interacted with any of the independent variables for any of the categorical response types. Furthermore, model comparisons showed that the full model was never significantly better than the same model with frequency removed as a predictor.

DURATION. The results for duration are reported in Table 9, and the means are in Figure 3. For this analysis, the independent variables were sequence type (VBV, VV, VGV) and rate (fast, slow). Duration was entered as a log-transformed variable. Another model was also run that included the identity of the initial vowel (V1) of the sequence (i, eɪ, o, u), which was fully crossed with the other two dependent variables. However, model comparison showed that the more complex model with V1 was not significantly better than the model that did not include that factor ($\chi^2(36) = 48.776$, $p = 0.08$). Likewise, another model that included frequency was also analyzed, but frequency and all of its interactions were not significant, and the model comparison showed that it was also not significantly better ($\chi^2(11) = 13.787$, $p = 0.245$). Thus, we report on the simpler model without V1 and frequency as factors.

The results show that there is a significant main effect of speech rate; the slow rate is significantly longer than the faster rate, which also confirms that participants spoke faster at that rate. (An examination of individual speakers shows that they all showed the same pattern of longer duration at the slower rate as compared to the faster rate.) There is also a significant main effect for VBV and VV compared to VGV; both of these have a significantly shorter duration. A Tukey post-hoc test collapsing over rate indicates that there is no significant difference in duration between VBV and VV sequences ($\beta = -19.07$, $z = -1.395$, $p = 0.34$). As shown in Fig. 3, the interaction between rate and VV is due to a smaller difference between this sequence and VGV in the fast condition than in the slow condition.

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	−0.283	0.070	−4.030*
seq: VV	−0.383	0.067	−5.680*
rate: slow	0.208	0.022	9.590*
seq: VBV * rate: slow	−0.029	0.027	−1.070
seq: VV * rate: slow	−0.065	0.020	−3.180*

TABLE 9. Statistical results for duration. In this and following tables, the baseline values for the independent variables are fast for rate and VGV for sequence type.

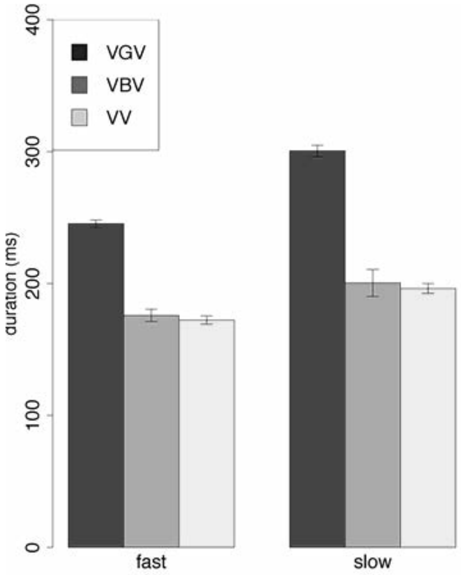

FIGURE 3. Duration of VGV, VBV, and VV sequences, grouped by rate.

Intensity. The analyses of intensity separately examined two variables—V1 intensity max − intensity min, and V2 intensity max − intensity min—in order to confirm that any effects of intensity are not dependent on which of these vowels is chosen for the analysis. The statistical results for intensity are reported in Table 10 and Table 11, and the means are in Figure 4. The dependent variables were sequence type (VBV, VV, VGV) and rate (fast, slow). A model was also run including glide (/w/ or /j/) as a factor that was fully crossed with sequence type and rate, but model comparison shows that there was no significant difference between the model that included the glides and the one that did not (V1 intensity: $\chi^2(18) = 10.63$, $p = 0.91$; V2 intensity: $\chi^2(11) = 17.33$, $p = 0.10$). Likewise, frequency did not contribute significantly to the model (V1 intensity: $\chi^2(11) = 5.43$, $p = 0.91$; V2 intensity: $\chi^2(11) = 5.65$, $p = 0.89$). Thus, the final model for both V1 intensity and V2 intensity matched that used for duration.

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	−2.4482	0.8240	−2.971*
seq: VV	−3.4960	0.7920	−4.414*
rate: slow	1.8361	0.3073	5.976*
seq: VBV * rate: slow	−0.2359	0.5150	−0.458
seq: VV * rate: slow	−1.5073	0.3899	−3.866*

Table 10. Statistical results for V1 intensity max – intensity min.

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	−3.784	0.713	−5.307*
seq: VV	−5.204	0.660	−7.891*
rate: slow	1.628	0.241	6.748*
seq: VBV * rate: slow	−1.161	0.395	−2.939*
seq: VV * rate: slow	−1.209	0.297	−4.079*

Table 11. Statistical results for V2 intensity max – intensity min.

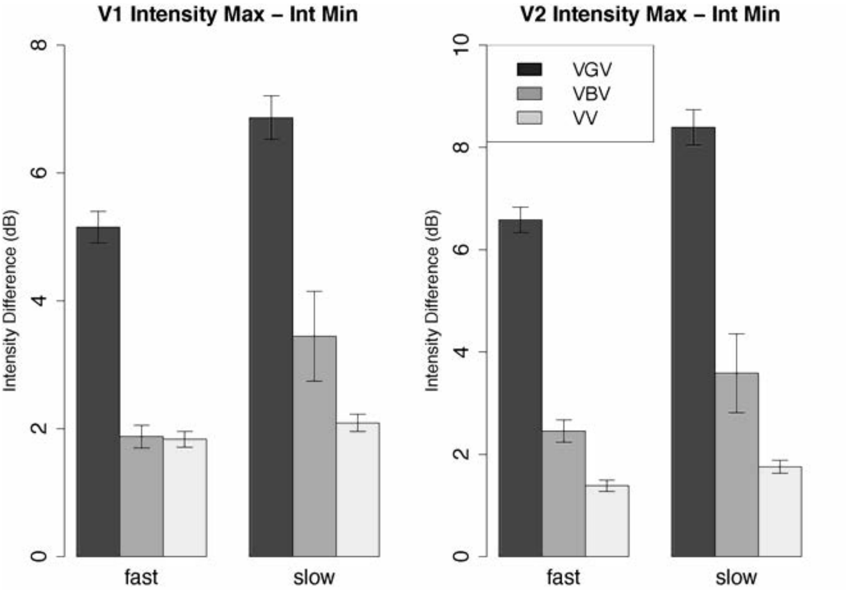

Figure 4. V1 and V2 intensity difference of VGV, VBV, and VV sequences, grouped by rate.

Results for V1 intensity max − intensity min show that the difference is significantly larger for VGV than for either VBV or VV. Intensity differences are also significantly larger at the slow rate than at the faster rate. The interaction between sequence type and rate is due to a difference between VBV and VV in the slow condition, but not in the fast condition. For V2 intensity max − intensity min, the intensity difference is also significantly larger for VGV than for VBV or VV. A Tukey post-hoc test collapsing over rate indicates that there is also a significant difference between VV and VBV ($\beta = -1.42$, $z = -2.70$, $p < 0.02$). The interactions arise because the intensity difference for VGV is higher in the slow rate than at the faster rate.

FORMANTS. For the analysis of the formants, /j/ and /w/ were analyzed separately. For /j/, a measure of F2 was taken, whereas F1 was measured for /w/. The statistical results are shown in Table 12 and Table 13, and the graph of the formant values is in Figure 5. For both formants, the independent variables were sequence type and rate. For /j/, frequency is also included as a variable. Though neither the main effect nor the interactions with frequency were significant in the final model, model comparison showed that the model without frequency was significantly different from the model that contained frequency ($\chi^2(11) = 20.74$, $p = 0.03$). For /w/, however, frequency was not a significant contributor ($\chi^2(11) = 10.22$, $p = 0.51$), so it was removed from the final model.

Results show that for /j/, F2 values are significantly higher for VGV as compared to both VBV and VV. Despite the seemingly large differences between VBV and VV in Figure 5, Tukey post-hoc tests indicate that there is no significant difference between VBV and VV, collapsing over rate ($\beta = -204$, $z = -1.92$, $p = 0.13$). Neither rate nor frequency are significant main effects, nor are any of the interactions significant.

For /w/, F1 values are significantly higher for VBV and VV as compared to VGV. The significant difference for the slow rate is attributable to VGV; F1 for VGV is significantly lower in the slow condition than in the faster speech. The interaction between VV and rate seems to be due to a bigger difference between VGV and VV in the slow condition than in the fast condition.[4]

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	−442.28	221.34	−1.998*
seq: VV	−700.30	288.88	−2.424
rate: slow	16.96	94.08	0.180
freq	−68.55	96.25	−0.712
seq: VBV * rate: slow	293.74	174.08	1.687
seq: VV * rate: slow	−17.47	110.35	−0.158
seq: VBV * freq	146.51	119.07	1.230
seq: VV * freq	−124.17	120.00	−1.035
rate: slow * freq	27.05	48.00	0.564
seq: VBV * rate: slow * freq	−97.20	68.43	−1.420
seq: VV * rate: slow * freq	10.69	62.38	0.171

TABLE 12. Statistical results for F2 values for /j/.

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	80.186	23.060	3.477*
seq: VV	75.249	21.506	3.499*
rate: slow	−23.592	5.584	−4.225*
seq: VBV * rate: slow	10.998	10.274	1.070
seq: VV * rate: slow	16.159	7.261	2.225*

TABLE 13. Statistical results for F1 values for /w/.

---

[4] If the reference value for sequence type is reordered so that VBV is the reference, then all of the significant differences for rate and the interactions disappear. This indicates that the significantly lower value of F1 in the slow condition for VGV than in the faster condition is causing both the main effect of rate and the interactions in Table 13.

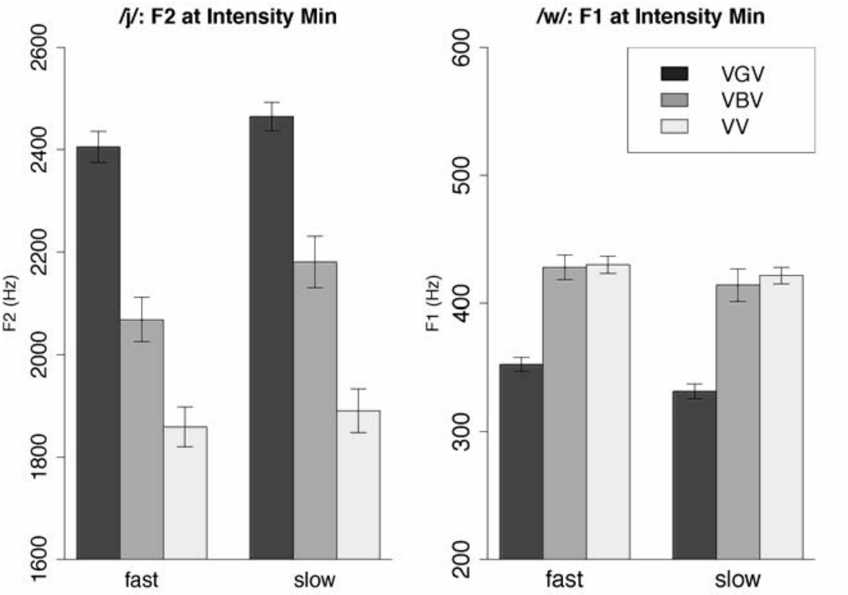

FIGURE 5. Formant values of VGV, VBV, and VV sequences, grouped by rate. The panel on the left is F2 in Hz for /j/, and that on the right is F1 in Hz for /w/.

CONTINUOUS VARIABLES FOR STRESSED V2 SUBSET. Following the results in §3.1, which showed that the production of glottal stops and glottalization were significantly higher when stress occurred on the second vowel of the VBV sequence (e.g. *see otters* [siʔɑɾɚz]), we also investigate whether there is any evidence that glides are produced before stressed vowels in these same environments for utterances coded as ‘modal’. It is possible that hiatus is resolved before stressed vowels, especially in VBV sequences. If so, this would produce formant values and intensities indicative of the presence of a glide before stress. To examine this question, we analyzed the formant values and intensity measures for the subset of VV and VBV tokens with stress on the second vowel as compared to the VGV tokens (all of which had lexical stress on the syllable containing the glide).

Results for both intensity calculations mirror the findings for the full data set. For both V1 intensity max − intensity min (Table 14) and V2 intensity max − intensity min (Table 15), there is a significantly smaller intensity difference for VBV and VV than for VGV. Intensity differences are also significantly larger at the slow rate than at the faster rate. The interactions between sequence type and rate arise because the intensity difference for VGV is higher in the slow rate than in the faster rate. Post-hoc Tukey tests for the main effect of sequence show that there are no significant differences between VBV and VV either for V1 intensity max − intensity min ($\beta = -0.559$, $z = -0.546$, $p = 0.84$) or for V2 intensity max − intensity min ($\beta = -1.45$, $z = -1.53$, $p = 0.28$).

	ESTIMATE	STD. ERROR	*t*-VALUE
seq: VBV	−3.108	0.926	−3.357*
seq: VV	−3.668	1.023	−3.584*
rate: slow	1.516	0.253	5.998*
seq: VBV * rate: slow	−0.140	0.583	−0.241
seq: VV * rate: slow	−1.201	0.523	−2.297*

TABLE 14. Statistical results for V1 intensity max – intensity min for sequences with stress on V2.

	ESTIMATE	STD. ERROR	t-VALUE
seq: VBV	−3.136	0.812	−3.862*
seq: VV	−4.582	0.911	−5.033*
rate: slow	1.638	0.279	5.864*
seq: VBV * rate: slow	−1.251	0.599	−2.088*
seq: VV * rate: slow	−1.361	0.535	−2.542*

Table 15. Statistical results for V2 intensity max – intensity min for sequences with stress on V2.

The same subset of the data was also used for an analysis of the differences in formant values—F2 for /j/ and F1 for /w/. Results for /j/ show that VGV has significantly higher F2 values than either VBV or VV (Table 16). There is no significant difference between VV and VBV ($\beta = -124.5$, $z = -0.702$, $p = 0.76$). For /w/, VGV has a significantly lower F1 value than either VBV or VV (Table 17), and there is no significant difference between VV and VBV ($\beta = -22.2$, $z = -0.806$, $p = 0.70$).

	ESTIMATE	STD. ERROR	t-VALUE
seq: VBV	−329.08	132.24	−2.488*
seq: VV	−453.59	165.63	−2.739*
rate: slow	62.50	31.37	1.992*
seq: VBV * rate: slow	−15.80	83.79	−0.189
seq: VV * rate: slow	−81.31	65.14	−1.248

Table 16. Statistical results for F2 values for /j/ for sequences with stress on V2.

	ESTIMATE	STD. ERROR	t-VALUE
seq: VBV	69.089	20.509	3.369*
seq: VV	91.300	27.189	3.358*
rate: slow	−23.363	5.751	−4.062*
seq: VBV * rate: slow	5.358	12.498	0.429
seq: VV * rate: slow	10.336	12.766	0.810

Table 17. Statistical results for F1 values for /w/ for sequences with stress on V2.

The results of the analyses of intensity and formants for the subset of data that contains only sequences with stress on the second vowel of the sequence indicate that VV and VBV are still significantly different from VGV, mirroring the pattern of the whole data set. Moreover, both for formants and intensity, there is no significant difference between VV and VBV, suggesting that it is not more likely for a glide to be produced in VBV than in VV.

Summary: continuous variables. The analyses of the continuous variables indicate the following results. First, though duration is only a very rough measure, VGV is significantly longer than VV or VBV. This finding is consistent with VGV sequences having three segments whereas modal VV and VBV sequences have only two. In addition, the significant difference between the fast and slow rates for the duration measure confirms that speakers were increasing their speech rate when told to read the passage faster.

The main result of the intensity analysis indicates that the difference between the maximum intensity of the vowels and minimum intensity of the VGV sequence—which corresponds to the glide—is significantly greater than the intensity difference for VBV and VV sequences. This can be seen visually in Figure 6, which shows a substantial dip in intensity for the glide in the VGV sequence for the /u + o/ combination that is not matched in degree by the VV or VBV sequences. This result is consistent with analyzing both VV and VBV as being a sequence of two vowels.

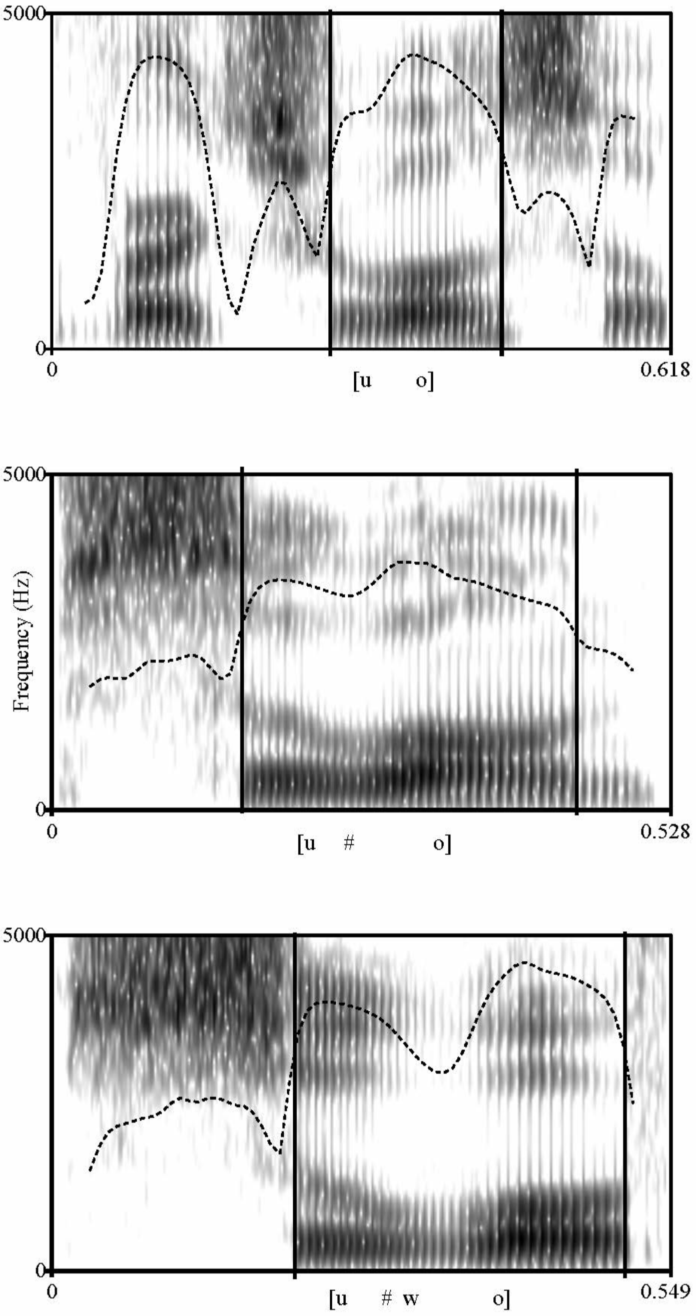

FIGURE 6. Intensity tracks (dashed lines) overlaid on VV *virtuoso* (top), VBV *Sue owned* (middle), and VGV *Sue wove* (bottom) sequences. The intensity range on the y-axis is from 35–70 dB.

For the measure V2 intensity max − intensity min for the whole data set, post-hoc tests also showed that there was a significantly greater intensity difference for VBV than for VV. This may be the result of differences in the articulatory coordination of vowel-vowel sequences within words versus across word boundaries (e.g. Browman &

Goldstein 1988, Byrd 1996, Hardcastle 1985). If there is less overlap of vowels in VBV sequences due to the intervening word boundary, then it is possible that more of the offglide of the first diphthongal vowel will be evident. If the offglide has a slightly lower intensity than the first part of the diphthong, this would account for the difference between VV and VBV sequences. It is worth repeating, however, that the intensity differences for both VV and VBV are much smaller than for VGV.

Results for the formant analysis indicate that the value of F2 taken at the intensity minimum for the sequences with a /j/ glide or high front offglide of V1 is significantly higher for VGV than for VV or VBV. Similarly for /w/, F1 is significantly lower for VGV than for the other two sequence types. These findings suggest that /w/ and /j/ in VGV sequences have maximally low and high target values, respectively, that are not matched by the vowel-vowel sequences. In other words, the formant results are not indicative of the insertion of a glide for either VV or VBV sequences. Although the differences in F2 for VV and VBV were not significant, they trended in a direction that is compatible with the hypothesis that there are articulatory timing differences between VV and VBV. The higher F2 values for VBV are consistent with a more robust offglide being produced for VBV than for VV.

Taken together, the results for the continuous variables demonstrate that for tokens produced with modal voicing, the acoustic properties of VV and VBV are very different from those of VGV, and are not consistent with what would be expected if there is glide insertion to resolve hiatus for vowel-vowel sequences. VGV is longer, has significantly lower intensity for the glide, and has more extreme formant values than VV or VBV. This is also true for the subset of data containing only stimuli with stress on the second vowel (e.g. *see otters*, *virtuoso*), indicating that there is no evidence of glide insertion to resolve hiatus even across word boundaries, which is the same environment that conditions greater amounts of glottal stop insertion before a stressed syllable. There is some evidence that there may be articulatory coordination differences between VV and VBV, such that the offglide of the first diphthongal vowel is more fully realized in VBV sequences, but this offglide still has very different properties from the glide in VGV sequence.

On the basis of these results, we argue that hiatus is not resolved with glide insertion in American English. The results do, however, lend themselves to a possible alternative interpretation: there is insertion of a glide-like element between the two vowels, but speakers attempt to minimize this epenthetic element in order to reduce the disruption to the underlying string of vowels. Yet another related possibility is that epenthetic elements may range from full glide insertion, to the insertion of a substantially reduced glide, to nothing at all, which would result in aggregated phonetic characteristics that are significantly different from a full glide. There may be some precedent for these alternative accounts. Hall (2013) shows that for some speakers, epenthetic [i] vowels in Lebanese Arabic differ from lexical vowels in the same segmental and stress environments, whereas other speakers show no significant differences in the acoustic properties of the two types of vowels. A related case is discussed by Yu (2007), who finds that tones in Cantonese lexical and morphologically derived words said to have the same rising pattern are actually not acoustically identical. Both Yu and Hall discuss how an exemplar-based model of categories could account for why there are acoustic differences in these cases, which they refer to as NEAR MERGERS.

While we cannot rule out the possibility that a reduced epenthetic element is being inserted to resolve the hiatus, without further evidence, the most parsimonious account of the data is that the percept of a glide-like element between the two vowels in the VV and VBV sequences is due to the diphthongal nature of American English vowels.

Moreover, it is interesting to note that in reports of languages that allow vowel hiatus (see examples in Casali 1998, 2011, though these are mainly about hiatus resolution), these languages are said to have monophthongal vowels. A future direction for this line of inquiry would be to examine whether languages that have a distinction between diphthongal and monophthongal vowels exhibit a three-way difference between sequences such as [e + V], [eɪ +V], and [e + jV]. While we leave room for the possibility that speakers are inserting a reduced or weak element, the remaining discussion and the phonological analysis that we present in §4 proceeds from our interpretation that English does not resolve hiatus via glide insertion.

4. General discussion.

4.1. Conditioning environments for hiatus (non)resolution. Taken together, the results from the categorical and continuous variables demonstrate that hiatus in American English has not been properly characterized in the phonological literature, which has simply assumed that hiatus is resolved with glide insertion after nonlow front and back vowels. Returning to the predictions in 2 in the introduction, the findings are generally consistent with 2d: hiatus in VBV can be resolved with a glottal stop (especially before a stressed vowel), but when a glottal stop is not produced, hiatus is allowed in VBV and VV sequences, and there are substantial acoustic differences between the VV and VGV sequences.

Although there is no apparent evidence for glide insertion, the pattern found in the current study is nevertheless surprisingly categorical. This is shown clearly in Figure 7, which collapses both the glottalization and glottal stop categories defined in §2.1 into a single glottal stop category, following the discussion in §3.1. While the modal responses for VGV and VV reach above 70%, the other main response is V2 creak, which is considered a voice-quality implementation and not an indicator of glottal stop insertion. For VBV, however, the single largest response type is glottal stop.

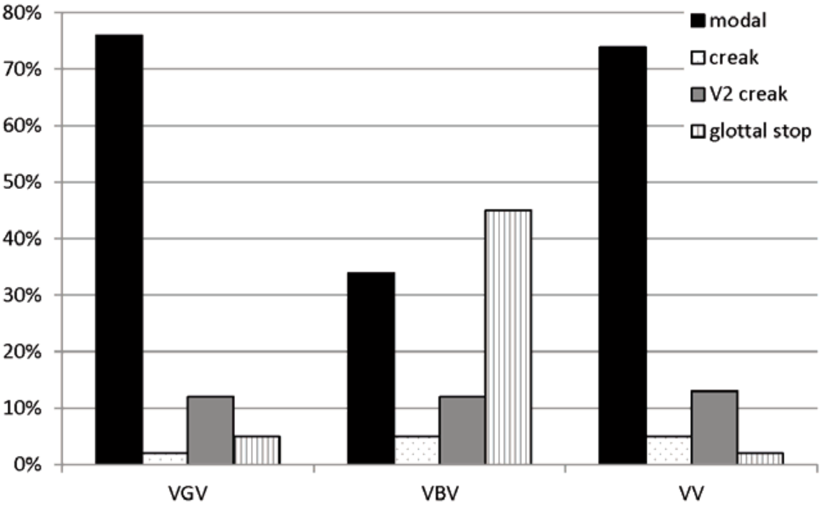

Figure 7. Percentages of response types for VGV, VBV, and VV sequences. 'Glottal stop' collapses over all of the data that were initially coded as either 'glottalization' or 'glottal stop'.

One question that arises based on the proportions in Fig. 7 is whether there are any factors that can account for why there is variation for VBV but very little for VGV and VV. This question can be addressed with the results from the analysis of stress in §3.1, which showed that there is variation between modal and glottal stop responses for VBV stimuli because glottal stop insertion across word boundaries is sensitive to the stress of the second vowel in the sequence (see Fig. 2). Although the data in this study were unbalanced for the variable of stress, a more balanced data set would likely confirm a clear overall pattern of hiatus resolution: little or no resolution in VV sequences and VBV sequences where V2 is unstressed, and substantial glottal stop insertion in VBV sequences where V2 is stressed. In the next section, we examine the phonological ramifications of these findings.

4.2. Phonological implications. In this section, we sketch out the main components of a phonological analysis of the results in an optimality-theoretic framework (OT; Prince & Smolensky 2004 [1993]). The resolution of hiatus in English depends on the interaction of two types of constraints: a family of Onset constraints that distinguish between different prosodic domains, and a constraint that prohibits nonlexical, or derived, glides in English. The Onset constraints that are most relevant for accounting for hiatus resolution are those proposed by Flack (2009), who defines a hierarchy of onsets that are sensitive to the prosodic domain. The relevant constraint schema is given in 3. The prosodic categories that Flack delineates are syllable, word, phrase, and utterance.

(3) M_{Ons} (Ons/PCat): Where M_{Ons} is some markedness constraint that targets onsets, and PCat is some prosodic domain, assign one violation for each instance of PCat whose leftmost onset violates M_{Ons}.

At first, for the word-boundary VBV sequences, it seems that a constraint requiring onsets at the level of the word—Ons/Wd—must outrank Dep. This would ensure that a segment is inserted before V2 in VBV sequences (e.g. *see [ʔ]otters*). The choice of [ʔ] as the epenthetic consonant can be attributed to its status as the least marked consonant for place (de Lacy 2006). As discussed in §4.1, however, hiatus resolution is not conditioned between all word boundaries, but rather occurs much more often when the first syllable of the second word is stressed. To capture this, the syllabic prosodic category must be divided into stressed and unstressed syllables, which Flack discusses. Because a constraint like Ons/σ́ would incorrectly pertain to both word-initial and word-medial stressed syllables, we also need a way of targeting only syllables that are both word-initial and stressed. One possible way to do so is to posit a conjoined constraint (Smolensky 2005) such as Ons/Wd & Ons/σ́, which would outrank Dep.[5] Ons/Wd & Ons/σ́ is violated by candidates in which the second vowel in an unresolved hiatus context is both word-initial and stressed. The tableau demonstrating glottal stop insertion before initial stressed syllables but not unstressed syllables is shown in 4 (the ranking Dep >> Ons/σ́ is not evident from this tableau, but it is established in 5). Although the constraints are presented with a fixed ranking in 4 for simplicity, a more complete phonological analysis based on a balanced data set would have to account for the variable results for glottal stop epenthesis at word boundaries (as demonstrated in Fig. 7).

(4)

s/i ɑ́/tters	Ons/Wd & Ons/σ́	Dep	Ons/Wd	Ons/σ́
s[i ɑ́]tters	*!		*	*
☞ s[i ʔɑ́]tters		*		
h/i ə/bjected				
☞ h[i ə]bjected			*	
h[i ʔə]bjected		*!		

---

[5] Since there have been arguments against local conjunction (e.g. McCarthy 1999, Padgett 2002), an analysis within harmonic grammar, which uses weighted constraints, may be a better solution (McCarthy & Pater 2015, Smolensky & Legendre 2005). In this article, however, we demonstrate the solution with the conjoined constraints for the sake of simplicity.

The second part of the pattern to be analyzed is the lack of glide insertion to resolve hiatus. Aside from the entrenched assumptions about English, glide insertion in which the glide is homorganic with an adjacent vowel has also been claimed to occur in languages such as Japanese (Kawahara 2003), Colloquial Slovak and Czech (Rubach 2000), and Polish (Rubach 2007), among others, though none of these claims have been verified acoustically. Thus, we posit a constraint on glide formation in hiatus contexts that could be ranked to either allow or disallow this phonological process. For the English data, the main intuition is that English does not have derived glides (Levi 2008); there are lexical vowels and lexical glides, but glides are not created by spreading or sharing vocalic features with a consonant. There have been numerous proposals in the literature that could be adapted to prevent this kind of spreading or sharing,[6] but we follow Rubach (2000) and adopt No-Multiple-Link (*Mult-Link) to ban candidates in which one feature is linked to more than one root node.[7] This constraint rules out candidates like *see [j]otters* and *pre[j]occupied*, where the epenthetic segment—the glide—is determined by the properties of the preceding vowel. The deterministic relationship between V1 and the glide indicates that the vocalic features of V1 spread onto a consonantal root node that has been epenthesized to prevent hiatus. The surface form *see [j]otters* would require both a violation of Dep and of *Mult-Link, though Dep (and *ʔ, shown in 5 for completeness) must be ranked low enough that glottal stop insertion is not prevented. Instead, the violation of *Mult-Link is the one that rules out forms with glide insertion.[8] The overall pattern for both VBV, where there is hiatus resolution mainly for stressed syllables, and VV, where there is no resolution, is shown in the tableaux in 5.

(5)

	Ons/Wd & Ons/σ́	*Mult-Link	Dep	Ons/Wd	Ons/σ́	*ʔ
s[i á]tters	*!			*	*	
s[i já]tters		*!	*			
☞ s[i ʔá]tters			*			*
☞ h[i ə]bjected				*		
h[i jə]bjected		*!	*			
h[i ʔə]bjected			*!			*
☞ pr[iá]ccupied					*	
pr[ijá]ccupied		*!	*			
pr[iʔá]ccupied			*!			*

---

[6] Related constraints that might be adapted to the hiatus-resolution case include CrispEdge (Itô & Mester 1999), UniqueAffiliation (Kawahara 2007), MultipleCorrespondence (Krämer 2008), and NoStraddling (Gouskova 2010).

[7] It is possible that *Mult-Link would have to be modified to refer to a feature that specifically pertains to glides; a broad ban on feature sharing may be too powerful for English, since there is some place assimilation between consonants, for example, as well as voicing assimilation. Such a feature might be [+vocalic], although there is not a clear consensus in the literature regarding the appropriate feature for differentiating glides from vowels (cf. Levi 2008, Nevins & Chitoran 2008, Padgett 2008).

[8] For the case of English, it might be possible to rule out glide insertion using a markedness constraint that prohibits inserting anything other than /ʔ/, which has been considered the least marked consonant for epenthesis (de Lacy 2006). This would obviate the need for a constraint like *Mult-Link, since one could just posit a ranking of *glide >> *ʔ. However, languages like Japanese, Czech, or Polish, which are said to have homorganic glide insertion, would still need a constraint like *Mult-Link to explain why the glide being inserted is the specifically homorganic glide (and that its features are derived from the preceding vowel). Otherwise, the insertion of the homorganic glide would be stipulative. Therefore, we use *Mult-Link here as a constraint that could account for the difference between English and languages like Japanese, Czech, or Polish.

In sum, according to our analysis, English resolves hiatus by glottal stop insertion and not gliding, and only at word boundaries before a stressed vowel. This is enforced by the high-ranked conjoined constraint Ons/Wd & Ons/σ́. Inserted glides are prohibited throughout (*Mult-Link >> Dep), and in medial contexts, hiatus is not resolved at all (Dep >> Ons/σ́).

Although this study attempts only to address the existence of glide insertion as a hiatus-resolution strategy in English, it is worth saying a few words about how these results bear on phonological analyses of /ɹ/-insertion as a hiatus-resolution strategy. Some analyses treat /ɹ/-insertion as analogous to the realization of glides after nonlow vowels: both processes are assumed to occur in order to satisfy an Onset requirement (e.g. Itô & Mester 2009, McMahon 2000, Uffman 2007). Since, however, the results of this study indicate that glide insertion does not occur in American English, this argument may no longer stand as a motivator for /ɹ/-insertion. But most /ɹ/-insertion environments discussed in the literature occur across word boundaries, so perhaps the proper analogy is with glottal stop insertion to satisfy Onset instead. It is, however, unclear whether the two environments are really parallel, since many of the /ɹ/-insertion examples in previous research include cases where the second vowel is unstressed or destressed, that is, *Shah[ɹ] of Persia*, *law[ɹ] and order*, *spa[ɹ] is* (Uffman 2007). In this study, an unstressed second vowel did not condition very much glottal stop insertion. Other /ɹ/-insertion environments that are discussed in the phonological literature occur within words, but at morphological boundaries, for example, *draw[ɹ]ing*, *withdraw[ɹ]al* (Itô & Mester 2009). Although we did not specifically control for morphological boundaries within the VV stimuli, there are some polymorphemic and compound words, such as *pr[iɑ]ccupied*, *panth[iɪ]sm*, *J[uɪ]sh*, and *l[eo]ver* (see Appendix A for other examples). Yet, as shown in Table 1, glottal stopping and glottalization accounted for only 2% of responses for all VV words, which indicates that glottal stopping as a hiatus-resolution strategy across morpheme boundaries simply cannot be comparable to /ɹ/-insertion in the same environments after low vowels. We leave it to future work to reconcile the phonological function of /ɹ/-insertion with the patterns of glottal stop insertion and faithful hiatus that occur in other environments in English.

5. Conclusion. The results of this study challenge long-standing, if somewhat anecdotal, assumptions about hiatus resolution in American English. The categorical and acoustic analyses provide no support for glide insertion as a hiatus-resolution strategy. Word-medially, hiatus is simply tolerated. At word boundaries, there is evidence that glottal stop insertion is the preferred strategy for hiatus resolution, but there is an interaction between resolution and stress, such that it occurs only before stressed syllables. Like word-medial position, hiatus is mostly allowed before unstressed vowels at word boundaries.

Despite the lack of any acoustic evidence for lexical glide insertion, it cannot be denied that over time, there has been such a strong percept of a glide that it has been taken for granted as a phonological process for English. This raises three related questions for future research. First, one reason that the percept of a glide may be so strong for English is because the relevant nonlow vowels are diphthongal. If that is the case, then what are the acoustic properties and the percept of vowel-vowel hiatus contexts for languages that have monophthongal vowels only? Relevant languages to examine for this might be Russian (Padgett 2008) or certain dialects of Spanish in which hiatus contexts coexist with diphthongal ones (e.g. Chitoran & Hualde 2007, Hualde & Prieto 2002). Second, in other languages that ostensibly have homorganic glide insertion, such as Japanese or Czech, would acoustic analyses confirm this phonological process? Even simple constraint rankings like those given above would predict that there are languages with actual homorganic glide insertion, so acoustic analyses of a number of relevant languages would be necessary to confirm that typological prediction. Finally, what are the acoustic characteristics and percepts of languages that do have hiatus, but that do not have all of the relevant lexical glides? In this case, certain vowel-vowel environments in Russian might be of interest, since Russian does have lexical /j/ but does not have lexical /w/ (Padgett 2008, Timberlake 2004). In cases where a vowel is preceded by /u/ or /o/ (e.g. examples in Gribanova 2008), is there any evidence of an intervocalic /w/? An examination of such languages, in comparison to English and similar languages, would shed light on whether the percept of a glide is possible for purely articulatory 'transitional' reasons (cf. Gick & Wilson 2006), especially if combined with an articulatory study.

Appendix A: Stimuli

VOWEL SEQUENCE	VBV (WORD BOUNDARY)	VGV (GLIDE AT WORD BOUNDARY)	VV (WITHIN WORD)
[i + (j)ɑ]	see otters (S)[9]	see yachts	kiosk
	he objected	see yonder	preoccupied (S)
[eɪ + (j)ɑ]	may honor (S)	day yacht	chaotic (S)
	café options (S)	pay yahoos	séance
[o + (w)ɑ/ə]	Echo Operations	snow wombat	Noah
	Joe Oscar (S)	shadow wasp	boa
[u + (w)ɑ]	too obvious (S)	two wandering	nuance
	new octopus (S)	few watchdog	skua
[i + (j)ɪ]	be ignorant (S)	be Yiddish	deities
	tea into (S)	see Yippee	pantheistic (S)
[eɪ + (j)ɪ]	pay immense	say Yiddish	Aramaic
	everyday important	obey Yin	archaic
[o + (w)ɪ]	know itchy (S)	go witness	stoic
	indigo ink (S)	no wizards	coincidentally
[u + (w)ɪ]	two images (S)	knew witches	pituitary
	tissue infected	knew whisky	Jewish
[i + (j)o]	Deepsea Ocean (S)	see yodelers	Theo
	she overheard	sesame yogurt	studio
[eɪ + (j)o]	gourmet oatmeal (S)	Sunday yoga	layover
	sauté okra (S)	bluejay yolks	Rodeo Drive
[o + (w)o]	so-so overnight (S)	Joe won't	coordinate (S)
	go over (S)	rainbow woven	co-owner (S)
[u + (w)o]	Sue owned (S)	Sue woke	duo
	two oboes (S)	Sue wove	virtuoso (S)

Appendix B: Reading passages

Story One. Last year our family took a unique summer vacation to a tropical island. My son Noah begged to see otters and boa constrictors, but the island was better known for its birds, especially the rare brown skua that lives in cliffs. Fortunately, he's interested in many areas of zoology so he wasn't too disappointed, especially when a hotel clerk said we might be able to see a rare creature called a snow wombat and a new octopus that the locals had spotted on the beach.

The rest of the family was excited to see yachts in the sea, because none of us had ever been on such a big boat before. One morning, two wandering boaters came to the hotel and offered us a trip on their day yacht, which was strangely called 'Séance for Skeletons'. One sailor said to us, 'Do you see yonder cliff? We will take you there.' But we turned them down, because it was too obvious that they were trying to overcharge us. Their pitch was not very nuanced and one sailor seemed very preoccupied with how much it would cost. A few watchdog websites had warned us about these kinds of scams. It was a bit chaotic when we asked them to leave, but finally they did.

Instead, we went to a ticket kiosk to find out about boat trips to a place called Shadow Wasp Island. The man in the booth gave us three choices. 'You may honor my family by sailing on our boat,' he said. 'The café

---

[9] (S) indicates that the first syllable of the second word is stressed.

options on our ship are very nice. Or there is a company called Echo Operations run by a good man named Joe Oscar who can take you there. Finally, you could pay yahoos from the countryside to take you there.' But he objected to that option because he said they don't know anything about the island. We chose to go on the yacht of the man's family. We had a wonderful afternoon seeing wildlife and learning about the island.

STORY TWO. When Libby was in college, she studied many religions. She read about the many deities of Hinduism, and about pantheistic Druid customs. She also learned that to obey yin and yang is important in some Chinese religions. Students were taught the principles of ancient languages like Aramaic in one class, and also how to say Yiddish greetings. She learned that her own name could even be Yiddish in origin, though she didn't think it was. She studied the stories of stoic leaders and unappreciated heroines trying to save their mother tongue. The professor of that class also told them to see 'Yippee: A Journey to Jewish Joy', a movie about Hasidic Jews in Ukraine.

Libby especially learned a lot about some archaic pagan practices that she studied in another class. In two images shown to the students, religious leaders were depicted carrying out both formal and everyday important tasks. She knew witches were responsible for leading many kinds of rituals, and learned that no wizards were allowed to wear white clothes, unless they wanted to pay immense fines. But, for example, she never knew whisky was thought to stop or delay infection. Nor did she know itchy lesions were treated with spells requiring a witch to throw tea into a boiling cauldron of indigo ink. Libby also learned that witches had coincidentally discovered that tissue infected with disease was usually treated with fluids from the pituitary gland.

By the end of the semester, Libby had broadened her horizons. She was glad she would no longer be ignorant about such interesting practices throughout history, and hoped in the future to go witness some of the things she had learned about in person.

STORY THREE. When she was growing up, Sue owned two oboes, a guitar, and a violin. She always knew she wanted to be a musician, and today she plays clarinet in a duo called Rainbow Woven Sky and oboe with a woodwind trio called Deepsea Ocean Blue. She was recently called a virtuoso by a prominent critic.

Last year, the trio went on tour. After a so-so overnight layover in Chicago, the group was on their way to Beverly Hills. The next morning Sue woke up after a great sleep, and went to a nearby diner for breakfast. She ordered the golden sesame yogurt to start. But she was still hungry so she also ordered the gourmet oatmeal with berries. As she was eating, she overheard some famous chefs talking. One said to the other, 'In a new recipe I sauté okra with bluejay yolks and ham.' Sue thought that sounded interesting, but they couldn't go to that restaurant because the bassoonist Joe won't eat meat.

To relax before their evening performance, the pianist Theo said he wanted to go shopping on LA's famous Rodeo Drive. Theo knew the co-owner of a shoe boutique and wanted to go over to say hi. Joe wanted to see yodelers perform in a nearby park, but Sue wanted to go to a studio called Sunday Yoga that a friend had recommended. The friends made a plan to coordinate again later. The trio played that night at a small concert hall. It was a wonderful performance, and the audience cheered. After the performance Sue wove in and out of the crowd, greeting her fans. The trip to California had been successful.

APPENDIX C: ILLUSTRATIONS OF CODING FOR CATEGORICAL VARIABLES

All of the following spectrograms come from the slow rate.

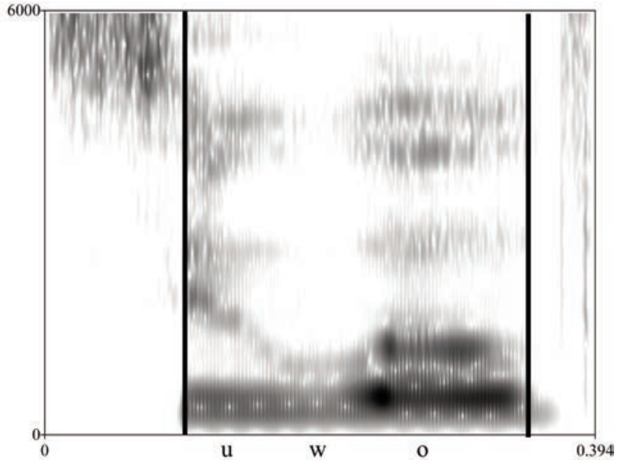

FIGURE A1. Modal response: VGV *Sue wove*.

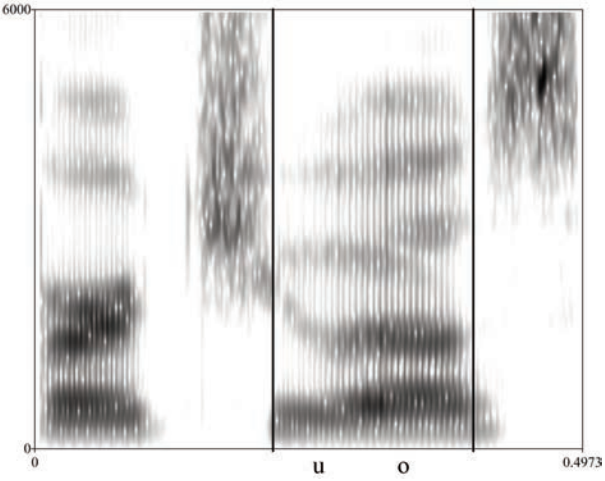

FIGURE A2. Modal response: VV *virtuoso*.

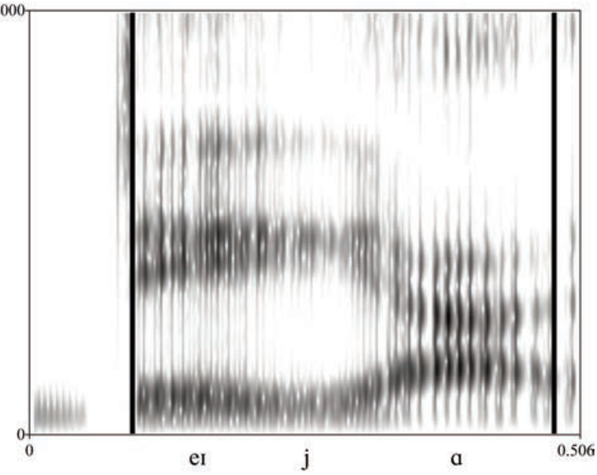

FIGURE A3. Global creak response: VGV *day yacht*.

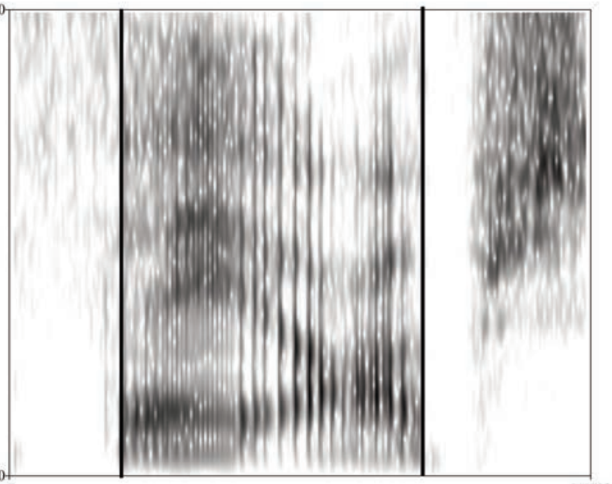

FIGURE A4. V2 creak response: VBV *café options*.

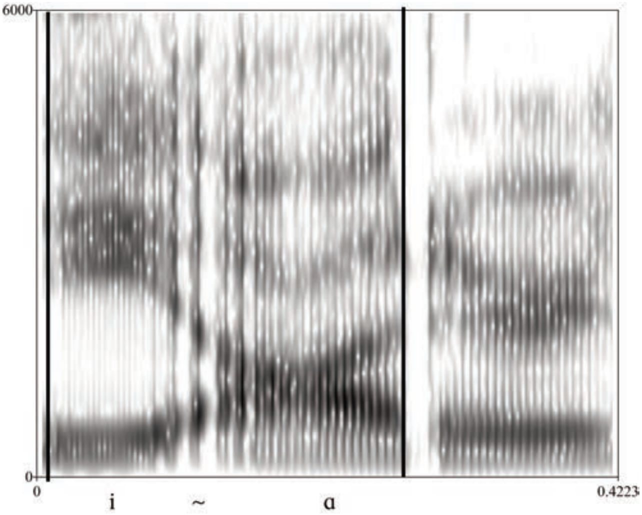

FIGURE A5. Glottalization response: VV *see otters*. The '~' between the vowels indicates a period of glottalized phonation.

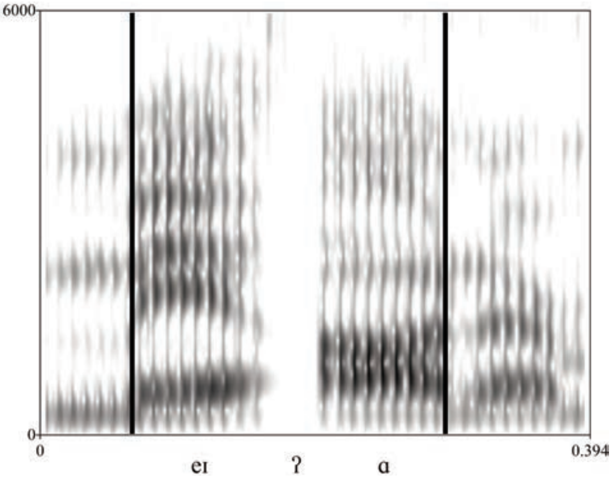

FIGURE A6. Glottal stop response: VBV *may honor*.

APPENDIX D: MULTINOMIAL ANALYSIS FOR CATEGORICAL RESPONSES

In this appendix, we model the categorical responses using multinomial logistic regression, which is a method that is appropriate when there are more than two possible responses for any given token. The analysis was carried out using the 'multinom' function in the nnet package in R (Venables & Ripley 2002). In this study, the VV, VBV, and VGV sequences were coded for modal, global creak, V2 creak, glottalization, and glottal stop responses (plus V1 creak and no response, which are not included in the multinomial analysis because there were too few such responses). In the binomial regressions reported in the main text, we compared each response type to all others, which were collapsed in each analysis. In a multinomial regression, a set of reference levels is also chosen, but the output of the analysis compares each of the other possible responses to the reference level. For this analysis, the reference levels are the modal response, the female speakers, the fast speech rate, and the VGV sequence. Since frequency was not significant in any of the binomial analyses, it is not included in the multinomial analysis.

In Table A1, the comparisons between each response type and the modal response are presented separately for ease of exposition, but it should be emphasized that these results are all from a single multinomial regression analysis. It should also be reiterated that this analysis does not include any random effects for either subjects or items, which will lead to a couple of different results from the binomial regression analyses reported in the main text.

MODAL VS. GLOBAL CREAK	ESTIMATE	STD. ERROR	$Pr(>\lvert z\rvert)$
(intercept)	−3.185	−9.627	0.000
gender: male	−1.527	−5.346	0.000*
rate: slow	−0.130	−0.523	0.601
seq: VBV	1.882	5.148	0.000*
seq: VV	1.077	3.013	0.003*
MODAL VS. V2 CREAK			
(intercept)	−1.418	−9.490	0.000
gender: male	−1.066	−7.053	0.000*
rate: slow	0.089	0.615	0.538
seq: VBV	0.816	4.448	0.000*
seq: VV	0.071	0.418	0.676
MODAL VS. GLOTTAL STOP			
(intercept)	−4.568	−11.351	0.000
gender: male	−0.376	−1.744	0.081
rate: slow	1.052	4.743	0.000*
seq: VBV	3.521	9.328	0.000*
seq: VV	−13.551	−0.044	0.965
MODAL VS. GLOTTALIZATION			
(intercept)	−3.121	−12.630	0.000
gender: male	−0.707	−4.188	0.000*
rate: slow	0.681	4.041	0.000*
seq: VBV	2.968	12.421	0.000*
seq: VV	−0.449	−1.323	0.186

TABLE A1. Multinomial analysis for categorical results.

The first point is a global one: the significant negative intercepts for each of the response types indicate that they are all significantly less likely to occur than the modal response. Since global creak and V2 creak have similar patterns of significance, they can be discussed in conjunction. First, male speakers are significantly less likely to produce either global creak or V2 creak, and there are no significant differences for rate, both of which are consistent with the findings for the binomial analyses. For the VBV sequence, speakers are significantly more likely to produce global or V2 creak than they are for VGV sequences. This result is consistent with the binomial analysis for global creak, but V2 creak was not significant in the binomial test. The VV sequence is significantly more likely to be produced with global creak than the VGV sequence is, but not with V2 creak. This is generally consistent with the binomial tests (for global creak in the binomial analysis, VV was marginally significant as compared to VGV at $p = 0.07$).

The results for glottal stop and glottalization can also be considered together, since they also have patterns of significance that are similar to one another. For gender, males are significantly less likely to produce glottalization than females; there is no significant difference for glottal stops. This finding is partially consistent with the binomial results, where gender was not a significant factor for either glottalization or glottal stops (though the directionality of the estimates was the same for both the binomial and multinomial analysis). The results for rate are the same for both response types: glottal stop and glottalization are both more likely to occur at the slow rate than the faster rate, as was the case with the binomial findings. As for sequence type, the VBV sequences are significantly more likely to have both glottal stops and glottalization than the VGV sequences are, and there are no significant differences between VV and VGV. (Though as noted in §3.1, the estimates for glottal stops are inappropriate because there are no glottal stops for VV, and only 0.004% of VGV sequences contain a glottal stop. But because of the way the multinomial analysis is carried out, we could not exclude this particular comparison from the test.) The results for sequence type match those for the binomial analysis.

The comparison between the binomial and multinomial analyses indicates that with two exceptions out of the many results provided by these analyses, the findings are consistent. Even where there are exceptions, the directionality of the estimates is the same and their magnitudes are similar.

REFERENCES

Aguilar, Lourdes. 1999. Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech Communication* 28.57–74.

Alber, Birgit. 2001. Regional variation at edges: Glottal stop epenthesis and dissimilation in Standard and Southern varieties of German. *Zeitschrift für Sprachwissenschaft* 20.3–41.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald; Douglas J. Davidson; and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59.390–412.

Baković, Eric. 1999. Deletion, insertion and symmetrical identity. *Harvard Working Papers in Linguistics* 7.1–11.

Balota, David; Melvin Yap; Michael Cortese; Keith Hutchison; Brett Kessler; Bjorn Loftis; James Neely; Douglas Nelson; Greg Simpson; and Rebecca Treiman. 2007. The English Lexicon Project. *Behavior Research Methods* 39.445–59.

Bates, Douglas M., and Deepayan Sarkar. 2008. lme4: Linear mixed-effects models using S4 classes. R package. Online: http://cran.r-project.org/web/packages/lme4/index.html.

Boersma, Paul, and David Weenink. 2011. Praat: Doing phonetics by computer. Version 5.2. Online: http://www.praat.org.

Borroff, Marianne. 2005. Articulatory phasing of glottal stop. *West Coast Conference on Formal Linguistics (WCCFL)* 24.70–78.

Bretz, Frank; Torsten Hothorn; and Peter Westfall. 2010. *Multiple comparisons using R*. Boca Raton, FL: Chapman & Hall/CRC.

Britain, David, and Susan Fox. 2008. Vernacular universals and the regularisation of hiatus resolution. *Essex Research Reports in Linguistics* 57.1–42.

Broadbent, Judith. 1991. Linking and intrusive r in English. *UCL Working Papers in Linguistics* 3.281–302.

Browman, Catherine, and Louis Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45.140–55.

Brysbaert, Marc, and Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41.977–90.

Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24.209–44.

Casali, Roderic. 1998. *Resolving hiatus*. New York: Garland.

Casali, Roderic. 2011. Hiatus resolution. *The Blackwell companion to phonology, vol. 3: Phonological processes*, ed. by Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, 1434–60. Malden, MA: Wiley-Blackwell.

Chitoran, Ioana, and José Ignacio Hualde. 2007. From hiatus to diphthong: The evolution of vowel sequences in Romance. *Phonology* 24.37–75.

Cruttenden, Alan. 2008. *Gimson's pronunciation of English*. 7th edn. London: Hodder Education.

de Lacy, Paul. 2006. *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press.

Dilley, Laura; Stefanie Shattuck-Hufnagel; and Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24.423–44.

Esling, John, and Jimmy Harris. 2005. States of the glottis: An articulatory phonetic model based on laryngoscopic observations. *A figure of speech: A festschrift for John Laver*, ed. by William Hardcastle and J. M. Beck, 347–83. Mahwah, NJ: Lawrence Erlbaum.

Espy-Wilson, Carol. 1992. Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English. *Journal of the Acoustical Society of America* 92.736–57.

FLACK, KATHRYN. 2009. Constraints on onsets and codas of words and phrases. *Phonology* 26.269–302.

GARELLEK, MARC. 2012a. Glottal stops before word-initial vowels in American English: Distribution and acoustic characteristics. *UCLA Working Papers in Phonetics* 110.1–23.

GARELLEK, MARC. 2012b. Word-initial glottalization and voice quality strengthening. *UCLA Working Papers in Linguistics* 111.92–122.

GELMAN, ANDREW, and JENNIFER HILL. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

GICK, BRYAN. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16.29–54.

GICK, BRYAN, and IAN WILSON. 2006. Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. *Laboratory phonology 8: Varieties of phonological competence*, ed. by Louis Goldstein, Douglas Whalen, and Catherine Best, 635–60. New York: Mouton de Gruyter.

GOUSKOVA, MARIA. 2010. The phonology of boundaries and secondary stress in Russian compounds. *The Linguistic Review* 27.387–448.

GRIBANOVA, VERA. 2008. Russian prefixes, prepositions and palatalization in stratal OT. *West Coast Conference on Formal Linguistics (WCCFL)* 26.217–25.

HALL, NANCY. 2013. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics* 41.133–43.

HARDCASTLE, WILLIAM. 1985. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication* 4.247–63.

HARRIS, JOHN. 1994. *English sound structure*. Oxford: Blackwell.

HESELWOOD, BARRY. 2006. Final schwa and r-sandhi in RP English. *Leeds Working Papers in Linguistics and Phonetics* 11.78–95.

HUALDE, JOSÉ IGNACIO, and MÓNICA PRIETO. 2002. On the diphthong/hiatus contrast in Spanish: Some experimental results. *Linguistics* 40.217–34.

HUNT, ELIZABETH HON. 2009. *Acoustic characterization of the glides /j/ and /w/ in American English*. Cambridge, MA: MIT dissertation.

ITÔ, JUNKO, and ARMIN MESTER. 1999. Realignment. *The prosody-morphology interface*, ed. by Rene Kager, Harry van der Hulst, and Wim Zonneveld, 188–217. Cambridge: Cambridge University Press.

ITÔ, JUNKO, and ARMIN MESTER. 2009. The onset of the prosodic word. *Phonological argumentation: Essays on evidence and motivation*, ed. by Steve Parker, 227–60. London: Equinox.

JAEGER, T. FLORIAN. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59.434–46.

JONGENBURGER, WILLY, and VINCENT VAN HEUVEN. 1991. The distribution of (word initial) glottal stops in Dutch. *Linguistics in the Netherlands* 1991.101–10.

KAWAHARA, SHIGETO. 2003. On a certain kind of hiatus resolution in Japanese. *Onin Kenkyuu* [Phonological Studies] 6.11–20.

KAWAHARA, SHIGETO. 2007. On the proper treatment of non-crisp edges. *Japanese/Korean linguistics 13*, ed. by Mutsuko Hudson Endo, Peter Sells, and Sun-Ah Jun, 55–67. Stanford, CA: CSLI Publications.

KOHLER, KLAUS J. 1994. Glottal stops and glottalization in German. *Phonetica* 51.38–51.

KRÄMER, MARTIN. 2008. English schwa insertion before liquids and phonological opacity. *Chicago Linguistic Society* 41.267–82.

LADEFOGED, PETER, and IAN MADDIESON. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

LEVI, SUSANNAH. 2008. Phonemic vs. derived glides. *Lingua* 118.1956–78.

MCCARTHY, JOHN J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38.169–95.

MCCARTHY, JOHN J. 1999. Sympathy and phonological opacity. *Phonology* 16.331–99.

MCCARTHY, JOHN J., and JOE PATER. 2015. *Harmonic grammar and harmonic serialism*. Sheffield: Equinox, to appear.

MCMAHON, APRIL. 2000. *Change, chance, and optimality*. Oxford: Oxford University Press.

Mompeán, Jose, and F. Alberto Gómez. 2011. Hiatus resolution strategies in non-rhotic English: The case of /r/-liaison. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, Hong Kong, 1414–17.

Nevins, Andrew, and Ioana Chitoran. 2008. Phonological representations and the variable patterning of glides. *Lingua* 118.1979–97.

Newton, Caroline, and Bill Wells. 2002. Between-word junctures in early multi-word speech. *Journal of Child Language* 29.275–99.

Ogden, Richard. 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31.139–52.

Orgun, Cemil Orhan. 2001. English *r*-insertion in optimality theory. *Natural Language and Linguistic Theory* 19.737–49.

Padgett, Jaye. 2002. Constraint conjunction versus grounded constraint subhierarchies in optimality theory. Santa Cruz: University of California, Santa Cruz, ms.

Padgett, Jaye. 2008. Glides, vowels, and features. *Lingua* 118.1937–55.

Pierrehumbert, Janet. 1995. Prosodic effects on glottal allophones. *Vocal fold physiology: Voice quality control*, ed. by Osamu Fujimura and Minoru Hirano, 39–60. San Diego: Singular.

Pierrehumbert, Janet, and Stefan Frisch. 1997. Synthesizing allophonic glottalization. *Progress in speech synthesis*, ed. by Jan P. H. van Santen, Richard W. Sproat, Joseph P. Olive, and Julia Hirschberg, 9–26. New York: Springer.

Pierrehumbert, Janet, and David Talkin. 1992. Lenition of /h/ and glottal stop. *Papers in laboratory phonology 2: Gesture, segment, prosody*, ed. by Gerard Docherty and D. Robert Ladd, 90–116. Cambridge: Cambridge University Press.

Prince, Alan, and Paul Smolensky. 2004 [1993]. *Optimality theory: Constraint interaction in generative grammar*. Cambridge, MA: MIT Press.

Quené, Hugo, and Huub van den Bergh. 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59.413–25.

R Development Core Team. 2012. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Online: http://www.r-project.org/.

Redi, Laura, and Stefanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29.407–29.

Rubach, Jerzy. 2000. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31.271–317.

Rubach, Jerzy. 2007. A conspiracy of gliding processes in Polish. *Journal of Slavic Linguistics* 15.325–57.

Smolensky, Paul. 2005. Optimality in phonology II: Markedness, feature domains, and local constraint conjunction. In Smolensky & Legendre, 27–160.

Smolensky, Paul, and Géraldine Legendre. 2005. *The harmonic mind: From neural computation to optimality-theoretic grammar*. Cambridge, MA: MIT Press.

Stene, Aasta. 1954. *Hiatus in English: Problems of concatenation and juncture*. Copenhagen: Rosenkilde and Bagger.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Uffman, Christian. 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29.451–76.

van Heuven, Vincent, and Annelies Hoos. 1991. Hiatus deletion, phonological rule or phonetic coarticulation? *Linguistics in the Netherlands* 1991.61–70.

Venables, William, and Brian Ripley. 2002. *Modern applied statistics with S*. New York: Springer.

Wiese, Richard. 1996. *The phonology of German*. Oxford: Clarendon.

Wolk, Leslie; Nassima Abdelli-Beruh; and Dianne Slavin. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice* 26.111–16.

Yu, Alan C. L. 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24.187–214.

Yuasa, Ikuko Patricia. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85.315–37.

Davidson
Department of Linguistics
10 Washington Place
New York, NY 10003
[lisa.davidson@nyu.edu]

Erker
Boston University
718 Commonwealth Avenue
Department of Romance Studies
Boston, MA 02215
[danny.erker@gmail.com]

[Received 26 June 2012;
revision invited 11 February 2013;
revision received 2 May 2013;
accepted 18 November 2013]