1 Introduction

Vowel insertion occurs in Turkish to repair both illicit complex codas and complex onsets. Complex codas that violate sonority sequencing occur in Arabic loanwords, such as /səbr/ 'patience.' These clusters are repaired by inserting a high vowel (1):

(1) Coda-repairing vowel insertion in Turkish
   a. /səbr/ [səbur] 'patience'
   b. /dʒebr/ [dʒeʃəɾ] 'algebra'
   c. /ømr/ [ømyɾ] 'life'

   The inserted vowel participates in vowel harmony, taking its backness and rounding from the preceding vowel (Clements & Sezer 1982). Like underlingly present vowels, the inserted vowel is stressed when it occurs in the final syllable. Since the Turkish coda-repairing vowel participates in the phonological processes of stress-assignment and vowel harmony, we can conclude that it is inserted during phonology – coda-repair is EPENTHESIS.

   Complex onsets in Turkish are also repaired by inserting a high vowel (2):

(2) Onset-repairing vowel insertion in Turkish
   a. /prəns/ [pɾəns] 'prince'
   b. /prəʋə/ [pɾəɾʋa] 'test'
   c. /bɾanda/ [bɾanda]¹ 'canvas'
   d. /bluʒin/ [buɾuʃin] ~ [byɾuʃin] 'blue jeans'

   Superficially, onset repair appears to be the same process as coda cluster repair, only in mirror image. Previous research (Yaş 1980, Clements & Sezer 1982, Kaun 1999, Yıldız 2010) characterizes both the onset- and coda-repairing vowels as epenthetic and harmonizing with the nearest root vowel. However, a closer look reveals real differences. The coda-repairing vowel is obligatory in speech and writing, while the onset-repairing vowel is optional in speech and absent in writing. Moreover, vowel-harmony is obligatory for the coda-repairing vowel, but...
variable and consonant-dependent for the onset-repairing vowel (Clements & Sezer 1982). These differences suggest that these two insertion processes are driven by different mechanisms: epenthesis for coda-repair, but INTRUSION for onset-repair. Intrusion, in contrast to epenthesis, occurs post-phonologically, at the level of articulation (Hall 2006). Consequently, intrusive vowels do not participate in phonology, and differ systematically from phonologically present vowels. Intrusive vowels are an acoustic percept resulting from gestural timing relations between consonants. Their presence is often optional or variable.

This paper argues that the Turkish onset-repairing vowel is intrusive, not epenthetic. Evidence comes from an ultrasound production experiment. Acoustic results indicates that onset-repairing vowel insertion is gradient and optional in Turkish, and that the inserted vowel is not a target for either backness or rounding harmony. Gesturally, underlying clusters produced with and without insertion turn out to be similar, despite their acoustic and perceptual differences, whereas perceptually similar inserted and underlying vowels display differing gestures. The specific gestural differences are consistent with the hypothesis that the inserted vowel lacks its own gestural target. These facts support an interpretation of Turkish onset-repair as a gestural, non-phonological phenomenon. This finding is significant for our understanding of Turkish syllable structure and vowel harmony, particularly since it conflicts with previous generalizations that were based on impressionistic data. This project also provides a methodological contribution through its use of ultrasound to probe a phonological question (à la Davidson & Stone 2003).

The paper is organized as follows: Section 2 describes the expected differences – particularly gestural differences – between intrusive and underlying vowels. Section 3 describes the ultrasound experiment designed to look for these differences between the onset-repairing vowel and underlying vowels. Results are presented in Section 4. Section 5 concludes.

2 Expected characteristics of intrusive vs. epenthetic vowels

Under a traditional division of phonology and phonetics, the phonological grammar cares about epenthesis, but is oblivious to intrusion. In the input to phonology, epenthetic vowels are absent; but in the output of phonology, epenthetic and underlying vowels are indistinguishable. The segments in the output of phonology map onto gestural targets (C, v, and r), whether v is epenthetic or underlying. The gestures produce an acoustic result, which the listener perceives
transarently as [Cvr].

Intrusion, on the other hand, creates an opaque relationship between the phonological output and the acoustic output. This occurs when the output of phonology remains [Cr] (no epenthesis), which in turn maps onto a series of two consonantal gestures. Depending on the exact timing relations of how those gestures are produced, the listener may perceive an intrusive vowel that was not present in the output of the speaker’s phonology. Table 1 schematizes the stages between phonological input and perceptual output for underlying, epenthetic, and intrusive vowels.

Table 1: /Cr/ vs. [Cvr]

<table>
<thead>
<tr>
<th>Input to phonology</th>
<th>Output of phonology</th>
<th>Perceptual result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying vowel</td>
<td>/Cvr/</td>
<td>[Cvr]</td>
</tr>
<tr>
<td>Epenthetic vowel</td>
<td>/Cr/</td>
<td></td>
</tr>
<tr>
<td>Intrusive vowel</td>
<td>[Cr]</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1

Gestural score for /Cr/ words

Since an intrusive vowel has no target, its tongue position will reflect the transition between the preceding consonant and the following vowel. Two gestures are involved, and their relative timing is schematized in Figure 1. In contrast, when the phonology yields [Cvr], there are three gestures, as shown in Figure 2. The underlying vowel $V_1$ is shaped by its own gestural target, not just its context. Thus, intrusive vowels are expected to differ gesturally from underlying vowels in ways that reflect the coarticulatory pressures of the preceding and following consonants and vowel, whose gestures overlap (Alfonso & Baer 1982, Öhman 1966). Note that the expected
gestural differences between perceptually-similar intrusive and underlying vowels vary, based on the relationship between the articulatory demands of $V_1$ and $V_2$.

Epenthetic and intrusive vowels might also differ in durational measures. In intrusion, the timing relation between the C and r gestures in a /Cr/ cluster is expected to vary over a smooth continuum between a close coordination that sounds like an insertionless cluster, and a more ‘pulled apart’ coordination that sounds like an inserted vowel. This timing relation should be reflected in the duration of the C_r interval, creating a monomodal distribution of C_r durations. With optional, categorical epenthesis, on the other hand, a bimodal distribution is expected, with one peak representing longer durations where epenthesis applied, and another peak representing short durations where epenthesis did not apply.

Finally, epenthetic and intrusive vowels are expected to pattern differently with respect to harmony. An epenthetic vowel should categorically participate (if harmony is root-governed, as suggested in Baković 2000) or fail to participate (if harmony is left-to-right, as traditionally argued (e.g. Underhill 1986)) in phonological vowel harmony. But an intrusive vowel can never be a target for phonological harmony, only for coarticulatory apparent harmony. Consequently, intrusive vowels might appear to participate in harmony in a gradient and variable fashion.

3 Method

To determine whether onset-repairing vowels in Turkish behave more like intrusive vowels, or more like epenthetic vowels, I conducted a fully factorial, 2 by 3 by 3 by 2 production experiment with six participants. The primary factor manipulated is the underlying structure of the target word: beginning with an onset cluster /Cr/, or beginning with a /Cvr/ sequence. To ensure that the findings extend across all consonant and vowel places, and investigate claims of vowel harmony in the inserted vowel, three stop consonants (/b/, /d/, /g/) and three vowels (/i/, /a/, /o/) were included. The final factor manipulated was the familiarity of the target word – either a real, familiar word, or a completely unfamiliar nonce word. This was to check that insertion is a fully productive process.

3.1 Materials

A list of real and nonce /Cr/ and /Cvr/ words was constructed (Table 2). Since multiple vowels
are attested as candidates for insertion before /i/ (both [i] and [ı]) and /o/ (both [u] and [ı]), there are two possible output [CVr] pronunciations for most input /Cr/ words, resulting in minimal triplets. I refer to minimal triplets and minimal pairs collectively as minimal sets. Within each minimal set, we controlled for: stress placement, the number of syllables, and the major place of articulation of C2. Stimuli are presented in Table 2, where: ‘-’ indicates a morpheme boundary, ‘.’ marks a syllable boundary, and ‘*’ indicates a word is being used as a /Cvr/ match for both the real and nonce /Cr/ words of that C-V condition. Numbers next to real /Cr/ words are familiarity ratings on a five point scale, where 5 is most familiar, obtained from three native speakers of Turkish.

Table 2: Stimuli

<table>
<thead>
<tr>
<th>C1</th>
<th>V2</th>
<th>Experimental</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Real /Cr/ word (familiarity) 'gloss'</td>
<td>Nonce /Cr/</td>
</tr>
<tr>
<td></td>
<td>/i/</td>
<td>bri. fin (4) 'briefing'</td>
<td>bri. mi. ti</td>
</tr>
<tr>
<td></td>
<td>/a/</td>
<td>bran. sı (4.33) 'subject.ACC'</td>
<td>brat. çi. ten</td>
</tr>
<tr>
<td></td>
<td>/o/</td>
<td>bro. sır (4.67) 'brochure'</td>
<td>bro. jör. le</td>
</tr>
<tr>
<td>d</td>
<td>/i/</td>
<td>drip. ling (1) 'dribbling'</td>
<td>drip. li. ke</td>
</tr>
<tr>
<td></td>
<td>/a/</td>
<td>dra. ma (4) 'drama' or dra. ma (3.7) 'drama.DAT'</td>
<td>dra. fa</td>
</tr>
<tr>
<td></td>
<td>/o/</td>
<td>bor. drom (4) 'payroll.my'</td>
<td>lor. dro. pur</td>
</tr>
<tr>
<td>g</td>
<td>/i/</td>
<td>grip (5) 'influenza'</td>
<td>gri. vi</td>
</tr>
<tr>
<td></td>
<td>/a/</td>
<td>gram (5) 'gram'</td>
<td>gra. bi</td>
</tr>
<tr>
<td></td>
<td>/o/</td>
<td>gro. s-u (2.67) 'gross.ACC'</td>
<td>gro. dol</td>
</tr>
</tbody>
</table>

Seventeen fillers were also included, primarily loanwords from English or French. Target words were presented in the carrier sentence in (3), which includes slots for two target words (X and Y). The sentence was designed to elicit contrastive focus on the target words.
To control for any effect of position within the sentence, half the repetitions employed an X-Y order, and the other half employed a Y-X order.

3.2 Participants

Seven native speakers of Turkish (4 female: S1,2,4,5,7) were recruited from the UC Santa Cruz community. Speaker 3 is bilingual in French and Turkish. The remaining speakers all studied English in school during adolescence, but lived in Turkey, using Turkish as their primary language at home and work, until age 18 or later, with the exception of one year spent in New Jersey for S6 (age 4-5). Participants were paid $20 for their time.

3.3 Data collection procedure

A consent form was provided in English. A language background questionnaire and experimental instructions were provided in Turkish. Participants were told that the purpose of the experiment was to study the way Turkish speakers' mouths and tongues move as they pronounce words.

Participants wore an Articulate Instruments Ultrasound Stabilization Headset (Wrench 2008) to stabilize the ultrasound probe. Recordings were made in a sound-attenuated booth using a shotgun microphone with a USB pre-amplifier connected to the ultrasound machine (Terason T3000 ultrasound system with a model 8MC3 probe). Subjects were asked to practice reading the instructions to get comfortable speaking with all the equipment, and were instructed to start the sentence over if they felt they had made a mistake. The experimenter also intervened when disfluencies or errors were noticed. Participants were requested to speak carefully and enunciate clearly, as if they were announcers on TRT (Turkish Radio and Television), whose broadcasters' careful articulation is famous in Turkey.

Stimuli were presented to participants on a laptop screen, with the target words already embedded in the carrier sentences. One sentence was visible at a time. Participants read through a list of 27 sentences (each containing up to two target words) five times. Within the list, all

\[2^{\text{S1}}\text{’s data are not discussed further because she participated in a pilot version of the experiment in which words began with /p t k/ and there were no fillers.}\]
sentences were randomized together. After each reading of the sentence-list, participants were offered the chance to take a break. At the end of the experiment, participants filled out a debriefing form with questions provided in Turkish as well as English. Responses indicated that participants understood the overall nature of the experiment (a linguistic investigation of the movements of the tongue) but not its specific purpose.

3.4 Data processing procedure

Acoustic annotation of the \( v_1 \) interval was conducted in Praat (Boersma & Weenink 2015) using TextGrids. The left edge of the \( v_1 \) interval was marked from the beginning of the \( C_1 \) release burst, identified by a dramatic increase in amplitude. The right edge of the \( v_1 \) interval was identified by the decrease in amplitude accompanying the onset of /r/.

One repetition of "drama" from S4 was not captured due to a recording error. Also, S2’s ultrasound data was discarded due to poor image quality (a result of individual anatomical features). As mentioned above, the condition \( C=/d/ \) and \( V_2=/o/ \) was also excluded from the analysis because the word-medial position of the cluster was a confound.

For the gestural analysis, I used a Python script to select the ultrasound frames best corresponding to beginning (onset), middle (midpoint) and end (offset) of the \( v_1 \) interval. This interval is quite short (generally 30-60 ms), so sometimes only one or two frames were captured. Tongue tracings were made in Edgetrak (Li et al. 2005), and SSANOVAs (see Davidson 2006) were fitted for each \( Cv_rV_2 \) combination, within subject. Since no significant acoustic differences were found between real and nonce words, the nonce/real distinction was collapsed.

4 Results

This section describes the distribution of durations; the rate of insertion and quality of the inserted vowels; and the gestural comparison of inserted and underlying vowels.

4.1 Distribution of durations

To determine whether insertion in Turkish onset clusters is a categorical or a gradient process, we looked at the duration of the interval between the release of the consonant and the beginning of the /r/ closure (the \( C_r \) interval). Density plots were made using R (R Core Team 2013) and the \texttt{density.compare} function in the \texttt{sm} package (Bowman & Azzalini 2014). All correctly recorded
tokens were plotted, including acoustically ambiguous tokens. In the following plots, the distribution of the duration of underlying vowels is also included for comparison.

The density plots show that underlying clusters, like underlying vowels, have monomodal distributions. This indicates that insertion is a gradient process, not a categorical one. However, the plot showing cluster data from all subjects (Figure 4) has a slight “hump” around 10ms, suggesting incipient bimodality. This goes away when S3, a bilingual, is excluded (Figure 3).

Figure 3: Duration distributions (all subjects)  Figure 4: Duration distribution (excluding S3)

Both Figure 4 and Figure 3 show that the C_r interval is shorter in underlying clusters than in underlying Cvr- words – the Cr distribution’s peak occurs about 20ms to the left of the Cvr distribution’s peak. That is, on average, the C_r is shorter in the underlying clusters than in the underlyingly Cvr words. This is predicted if there is no phonological epenthesis, and consequently no gestural target or timing slot between C and r in the underlying clusters.

4.2 Rate of insertion and quality of inserted vowels

To determine when insertion had occurred and what the quality of the inserted vowel was, we solicited perceptual judgements of /Cr/ word from nine phonetically-trained graduate students in the linguistics program at UC Santa Cruz. Judgements were collected in a SuperLab 4.5 experiment conducted in a sound-attenuated booth. Coders were asked to decide whether each production of a consonant cluster contained an inserted schwa, /i/, /u/, other vowel, or no vowel.

3For three tokens, participants provided judgments in a separate session outside of the laboratory setting, because the files included in the experiment contained errors.

4Schwa was included as a response option in place of [ı] for two reasons: first, because the coders were English
Jennifer Bellik

(faithful production with no insertion). These response options were selected on the basis of previous studies of onset repair.

Participants reported that coding the inserted vowels was surprisingly difficult. This was probably due to the short duration of the vowels (50-70 ms), and the fact that participants were not Turkish speakers. The difficulty of the task was reflected in the low degree of agreement among annotators, with all 9 participants agreeing for only 17 tokens (3.5% of the data). Interannotator reliability was fair (Fleiss' kappa = 0.228, Krippendorf's alpha = 0.228). However, this is still much better agreement than chance, considering that 5 responses were available.

Since interannotator agreement was so low, the remainder of the analysis excludes tokens that were coded with less than 90% confidence. Confidence was calculated according to the formula in (4), derived from Bayes' theorem.

\[
(4) \quad \text{Confidence} = P(sV = V | R)
\]

where \(R\) = set of responses, \(V\) = actual vowel, \(sV\) = vowel selected by all participants together (‘mode’), and throughout the probability calculations, \(sV\) is used as an approximation for \(V\).

When confidence was above 90% (74% of the dataset), the token was coded as containing the mode of the nine coders' responses, with the exception that tokens perceived to contain inserted \([u]\) before unrounded vowels were re-coded as containing \([i]\). This was because rounding of \(v_1\) before an unrounded \(V_2\) was entirely unprecedented in previous literature, and English speakers are known to be very unreliable at distinguishing rounded and unrounded back vowels (Lisker 1989). Tokens for which confidence was below 90% or for which the majority response was “other” were excluded from analysis, with the exception of tokens for which coders were divided between \([u]\) and \([i]\) responses.

Among /Cr/ tokens coded with >90% confidence, inserted vowels were perceived about half the time (157 / 352 Cr tokens = 44.6%), and most of the inserted vowels were perceived as schwa or \(<i>^5\) (108/157 = 68.8%). Consistent with previous studies and the corpus results, schwa

speakers who have schwa but not /u/ as a phoneme, and second, because Hall (2006) characterizes of the quality of intrusive vowels cross-linguistically as “either schwa, a copy of a nearby vowel or influenced by the place of the surrounding consonants.”

\(^5\)Inserted vowels are transcribed between angle brackets, and underlying vowels between square brackets: \(<i>^5\) refers to an inserted vowel perceived as /i/, and \([i]\) refers to the surface form of an underlying /i/.
was found inserted before any main vowel, but inserted <i> only occurred before main vowel /i/.

Figure 5: Insertion rate/quality for all speakers

![Vowel insertion (all speakers)](image)

Palatal harmony (insertion of <i> before /i/, shown in green) and rounding harmony (insertion of <u> before /o/, shown in darker blue) very rarely obtain. These results conflict strikingly with previous reports of regressive harmony in onset repair.

4.3 Gestural results

We now turn to the gestural analysis. In the following discussion, I refer to a \( CV_1 V_2 \) combination as a WORDTYPE. In this experiment, there were 15 potential wordtypes, of which nine were produced with insertion enough times to create an SSANOVA for at least one speaker. Only results for the midpoint are reported; generally, onsets and offsets look quite similar.

4.3.1 /b/ conditions

As a labial consonant, /b/ does not have an impact on tongue gestures, so the gestural effect of the following vowel should be particularly clear in the /b/ conditions. Table 3 recapitulates the /b/ conditions in which enough insertion occurred for an analysis for at least one speaker. Cells representing unexpected and unattested wordtypes are greyed out.
Table 3: /b/ conditions with analyses

<table>
<thead>
<tr>
<th>inserted v</th>
<th>b_ri</th>
<th>b_ra</th>
<th>b_ro</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;i&gt;</td>
<td>S5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;u&gt;</td>
<td>S7</td>
<td>–</td>
<td>S4</td>
</tr>
</tbody>
</table>

[bir]. For this wordtype, gestural results come from S5 (Figure 6). In the inserted vowel, the tongue is lower and backer compared to its position in the underlying /i/. The underlying /i/ has already attained its [+high, +front] target, but during production of the intrusive <i>, the tongue is still approaching the target associated with the following /i/. Fronting and raising only begin at the release of the consonant (Alfonso & Baer 1982), so the tongue is unlikely to attain the gestural target for the following /i/ during <i> in the C_r interval.

Figure 6

[bi]. Gestural results in this condition come from S7 (Figure 7), and show the tongue being lower and fronter in the /bri/ words (with and without insertion) than in /biri/. The intrusive hypothesis predicts this: the tongue is preparing to produce the following vowel /i/ during the b_r interval, but is unlikely to attain the /i/ target before /r/, because fronting and raising only begin at the release of the consonant /b/, the onset of the v1 interval (Alfonso & Baer 1982). In contrast, in /biri/ words with an underlying vowel, the tongue needs to achieve a [+high, +back] target for the underlying /u/ during the C_r interval. Therefore, tongue position should be fronter and lower in intrusive [i] than in underlying [i]. These predictions are borne out.
[buro]. Figure 8 shows the SSANOVAs for S4, the only subject who produced enough [b<u>ro] tokens for an analysis. The tongue is higher and fronter in underlying /br/ words produced with and without a cluster than in words with an underlying /u/. Thus, the inserted vowel is articulated in the same way as the cluster, but quite differently from the underlying vowel. This is evidence that the inserted vowel is intrusive.

When /bro/ is produced as [b<u>ro], intrusive <u>’s tongue position might be expected to reflect anticipatory coarticulation for the following mid vowel /o/. This lowering should begin at the offset of the preceding consonant (Alfonso & Baer 1982) – i.e. the beginning of the b_r interval. The underlying [u] in /buro/, on the other hand, has a [+high] target to hit during the b_r
interval. This predicts an intrusive <u> will be lower than the underlying [u] – the opposite of S4’s data.

Instead, it appears that S4 may be coordinating lowering the offset of /r/, rather than the offset of /b/. This would mean that no lowering would be apparent during the b_r interval. Instead, the tongue could still appear high, following the jaw closure necessary to produce /b/. The underlying /u/, on the other hand, would require a lower tongue position, because even a high vowel requires relatively unobstructed airflow.

Summing up, in all the /b/ conditions, there are significant gestural differences between inserted and underlying vowels. Meanwhile, inserted vowels and insertionless clusters do not differ significantly. This is evidence that onset-cluster repair is a matter of timing and intrusion, not the phonological addition of a gestural target.

4.3.2 /d/ conditions

In /d/ conditions, the preceding /d/ is expected to have a weak fronting effect on both underlying vowel and intrusive vowels, due to tongue coupling. Consequently, gestural differences between intrusive and underlying vowels will be less clear in /d/ conditions than in the /b/ conditions, where the labial consonant does not directly affect tongue position and so coarticulation with the following vowel is expected to begin earlier. Wordtypes with data are presented in Table 4.

Table 4: /d/ conditions with analyses

<table>
<thead>
<tr>
<th>inserted v</th>
<th>d_ri</th>
<th>d_ra</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;i&gt;</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>&lt;ı&gt;</td>
<td>S4, S7</td>
<td>S4, S5, S7</td>
</tr>
<tr>
<td>&lt;u&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[drı]. When /drı/ is pronounced with an intrusive <ı>, the tongue’s position will be determined by the transition between [+front] /d/ and [+front] /ı/. Fronting in preparation for /ı/ will begin at the /d/’s release (or potentially at the /r/’s release), so the tongue is unlikely to achieve /ı/’s [+front] target during the C_r interval. For underlying /ı/, the tongue needs to hit a [+back] target during the C_r interval. This predicts a fronter tongue body in the inserted vowels than in the underlying vowels.

There is gestural data from S4 and S7 for this condition. In accordance with the prediction
that fronting in <ı> begins when /d/ is released, in S4’s data (Figure 9), the tongue is significantly fronter in the inserted vowels. In S7’s data (Figure 10), the three tongue contours for [dr], [d<ı>r] and /dir/ all overlap – suprisingly, given that [dr] and /dir/ are so different perceptually. The dip in S7’s tongue in the inserted vowels probably represents the onset of the articulation of /r/. These results are consistent with the intrusive hypothesis, since the inserted vowels appear more coarticulated with /ı/ (for S4) and with /r/ (S7) than the underlying vowels do.

[dıra]. When /dra/ is produced as [d<ı>ra], if the <ı> is intrusive, then it reflects the transition between /d/ and the following low vowel /a/. Due to the influence of /a/, we expect the targetless <ı> to have a lower tongue body than the underlying /ı/ with its [+high] target. This prediction is borne out in the SSANOVA as for S4 (Figure 11) and S5 (Figure 12), where inserted vowels and insertionless clusters are lower than underlying vowels. But for S7 (Figure 13), there are no significant differences between clusters, inserted vowels, and underlying vowels. Perhaps this speaker reduces the underlying /ı/ in this context, or produces it at the very low end of its articulatory range.

To sum up, in the /d/ conditions, S4 and S5’s gestural results display the anticipatory coarticulation we predicted, while S7’s gestural results instead show that it is possible for clusters and underlying vowels to have the same gestural trajectory, despite their differing acoustics.
4.3.3 /g/ conditions

As a velar consonant, /g/ can be expected to contribute backing and/or raising to an adjacent vowel (Padgett 2011); this occurs, for example, in intrusive vowels in Maxakalí (Gudschinsky et al. 1970, Clements 1991, cited in Padgett 2011). However, as is common cross-linguistically, Turkish velar consonants tend to palatalize in the context of an adjacent front vowel — that is, the place of the vowel dominates the place of a velar consonant (Göksel & Kerslake 2005).
[gri]. When /gri/ is produced as [g<ı>ri], during the intrusive <ı>, the tongue is moving forward in preparation for /i/. This fronting could even begin during articulation of /g/, resulting in a palatalized /g/: fronting only begins at the release in /pVp/ sequences (Alfonso & Baer 1982), but could plausibly begin earlier with dorsal consonants than labial consonants. In underlying /ı/, the [+back] target of underlying /ı/ separates /g/ from /ı/, so /g/ would not palatalize (Göksel & Kerslake 2005), and fronting even during the C_r interval will be limited. Therefore, the intrusive <ı> is predicted to be fronter and higher than underlying /ı/. The difference in backness between underlying and inserted [ı] in this context should be clearer here after /g/ than it is after /d/, since /d/ has a confounding fronting effect.

The prediction that the inserted vowel will be fronter is borne out clearly for S6 (Figure 15). But for S4 (Figure 14), S5 (Figure 16) and S7 (Figure 17), there does not seem to be any difference in articulation between inserted and underlying vowels. This may reflect a lack of gestural differences between even the insertionless clusters and underlying vowels, as seen in some of the /d/ conditions above.

Table 5: /g/ conditions with analyses

<table>
<thead>
<tr>
<th>inserted v</th>
<th>g_ri</th>
<th>g_ra</th>
<th>g_ro</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;ı&gt;</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;ı&gt;</td>
<td>S5, S7</td>
<td>S4, S5, S6, S7</td>
<td>S4, S6</td>
</tr>
<tr>
<td>&lt;u&gt;</td>
<td></td>
<td></td>
<td>S3, S4, S6</td>
</tr>
</tbody>
</table>

Figure 14

Figure 15

Speaker 4

Speaker 6

Speaker 4

Speaker 6
[gra]. When /gra/ is produced as [g<ı>ra], the tongue begins lowering in anticipation of the following low vowel /a/ at the onset of the syllable (/g/). At the same time, /g/ tends to have a raising effect on adjacent vowels, which may counteract the lowering of /a/. Meanwhile, in /gira/ words with an underlying /ı/, both the raising from /g/ and the [+high] target of /ı/ will block this lowering.

Gestural data in this condition comes from S4 (Figure 18), S5, (Figure 19), S6 (Figure 20) and S7 (Figure 21). The primary gestural difference is that the underlying vowel appears backer and perhaps higher, particularly in the dorsum. This suggests that /a/ has a centralizing as well as

---

6The onset is included instead of the midpoint for this speaker/condition because of technical problems related to the smoothing factor in the SSANOVA for the midpoint.
a lowering effect. This counteracts /g/’s raising and backing effect in the intrusive vowels, while in the underlying vowels, the [+high, +back] targets of /g/ and /ı/ resist /a/’s influence. The difference is especially clear for S6 (Figure 20).

[gıro]. When /gro/ is produced as [gıro], the intrusive <ı> is likewise expected to have a lower tongue position than the underlying /ı/ from /gıro/, due to the lowering effect of the following /o/. Underlying /ı/ has a [+high] target which will make it more resistant to lowering.

[gıro]. For this condition, we have gestural data for S3 (Figure 23), S4 (Figure 25) and S6

In the gestural data for S4, the inserted vowel is indeed lower than the underlying vowel. In the gestural data for S6, on the other hand, the inserted vowel is no lower than the underlying vowel – but the no-insertion cluster is equally high, too.

[gıro]. For this condition, we have gestural data for S3 (Figure 23), S4 (Figure 25) and S6
Figure 26. Results are similar to those in the [gíro] condition, with height differences being subtle but significant, and accompanied by differences in backness. For all subjects, at least part of the tongue is lower in the inserted <u>, as predicted by coarticulation with /o/. For S4, the underlying vowel is also backer than the inserted vowel, similar to the [gíra] condition. Possibly the lowering induced by /o/ has a centralizing effect on the tongue. This is consistent with the intrusive hypothesis, since an intrusive <u> is expected to be lower than the underlying /u/ in /guro/, because the intrusive vowel will reflect the transition to low vowel /o/. Meanwhile, underlying /u/ adds a [+high, +back] target between /g/ and /o/, blocking /o/’s lowering effect.

To sum up, results in the /g/ conditions were somewhat harder to interpret than in the /b/ and /d/ conditions. Nonetheless, in every /g/ condition, most subjects show small but significant
gestural differences between inserted and underlying vowels – differences that can be interpreted to support the intrusive hypothesis.

4.4 Summary and discussion

The following table summarizes the gestural differences found in this study.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gestural findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>biri</td>
<td>Inserted vowel is lower/backer than underlying vowel (S5)</td>
</tr>
<tr>
<td>biri</td>
<td>Inserted vowel is fronter (S7)</td>
</tr>
<tr>
<td>buro</td>
<td>Inserted vowel is higher (S4).</td>
</tr>
<tr>
<td>dıri</td>
<td>Inserted vowel is fronter/lower (S4, S7)</td>
</tr>
<tr>
<td>dıra</td>
<td>Inserted vowel is lower for S4 and S5, but not different for S7.</td>
</tr>
<tr>
<td>gıri</td>
<td>Inserted vowel is fronter for S6 and maybe S4, but not for S5 or S7.</td>
</tr>
<tr>
<td>gıra</td>
<td>Inserted vowels seem more central and perhaps lower (S4, S5, S6, S7)</td>
</tr>
<tr>
<td>gıro</td>
<td>Inserted vowel is lower for S4, but not S6.</td>
</tr>
<tr>
<td>guro</td>
<td>Inserted vowel for S3, S4 and S6 is lower and more central.</td>
</tr>
</tbody>
</table>

This study found significant gestural differences between inserted vowels and underlying vowels. In some conditions, these differences were in the direction predicted by a priori phonetic expectations, and for some subjects, the differences were extremely clear. In other conditions, it was less clear how to interpret the differences, or the differences were more subtle, or there were no significant gestural differences between insertionless clusters, inserted vowels, and underlying vowels. But in every condition, the articulation of the inserted vowel by at least one subject (and usually by most subjects) was influenced by the following vowel in a way that the underlying vowel was not. This supports the intrusive hypothesis that the inserted vowels are targetless and therefore helpless to resist the coarticulatory effects of the preceding consonant and following vowel. The gestural differences also suggest that the initial consonant in the consonant cluster is tautosyllabic with the following lexically present vowel, since movement in preparation for that vowel usually begins at the start of the C_r interval. In short, the gestural differences, while not entirely clear-cut, support the intrusive hypothesis.


5 Discussion and conclusion

This study was concerned with two phonetic/phonological processes: vowel insertion and harmony (on the inserted vowel). Both processes are found to be variable/gradient. Perceived insertion occurs some of the time (~45%). Apparent harmony occurs some of the time (a much lower percent). Consequently, if insertion and harmony in onset-cluster repair are phonological, they must be optional. Also, inserted vowels differ from underlying vowels in their duration (they are shorter). Thus, if insertion and harmony are phonological processes, they must be gradient, in addition to being optional. That is, insertion must insert a vowel that is "less" of a vowel than the underlyingly present vowels – or must "partially insert" the vowel. If we allow phonological processes that are both optional and gradient, we can still conceive of Turkish onset-cluster repair as epenthesis. But with so much variability, it is simpler to think of onset-cluster repair as a phonetic process of intrusion, which would inherently be gradient.

If the onset-repairing vowel is intrusive, as argued here, then its behavior should not be used as the basis for arguments about vowel harmony and syllable-structure in Turkish phonology. In regards to harmony, an intrusive vowel’s harmonic behavior does not bear on questions about harmony as a phonological process. An intrusive vowel has no phonological presence and therefore cannot be a target for phonological harmony. This suggests that the reasoning behind studies where the behavior of the onset-repairing vowel is used as a basis for claims about phonological harmony (e.g. Kaun 1999) – must be re-evaluated. In addition, the non-harmonizing behavior of the inserted vowel cannot be used as an argument to bolster the traditional understanding of vowel-harmony in Turkish as being a strictly left-to-right process, since an intrusive vowel could never be a target for phonological harmony anyway; neither can its occasional harmonic appearance – actually due to coarticulation – be attributed to the emergence of a normally invisible right-to-left harmony pattern. This study does speak to questions about the phonetic and gestural origins for phonological harmony, providing some examples of anticipatory coarticulation creating the percept of harmony (the occasional instance of <i> insertion before /i/, or the more frequent instances of <u> insertion before /o/).

Additionally, with regard to syllable structure, since the onset-repairing vowel is not phonologically present, we can conclude that there is no categorical prohibition of complex
onsets in the foreign stratum of Turkish phonology. Rather, gestural timing relations create the percept of a vowel in a sequence that, phonologically speaking, remains a complex onset. Of course, there is always the possibility for hearers to reinterpret the acoustically ambiguous inserted vowel as representing a phonologically present vowel; this is probably the source of orthographic alternations like stil ~ sitil 'style'.

This project also bears on the extensibility of Davidson & Stone (2003)'s methodology to other phonological problems. This study applies D&S's experimental design but combines it with a more modern statistical technique, SSANOVAs, for a more nuanced analysis. Overall, I find that this comparative ultrasound methodology was successful in probing the phonological status of Turkish onset-cluster repair, but that the gestural results were best interpreted in conjunction with an analysis of the duration of the inserted vowels, since the differences between inserted and underlying vowels were often subtle, and varied significantly from subject to subject. Indeed, the gestural findings from this study suggest that there can be a great deal of interspeaker variability in the articulation of sequences that are not contrastive in a language. Further research with more speakers could illuminate this issue.

REFERENCES


