JOURNAL ARTICLE

An acoustic study of vowel intrusion in Turkish onset clusters

Jennifer Bellik

Department of Linguistics, University of California, Santa Cruz, CA, US

jbellik@ucsc.edu

Studies in Articulatory Phonology (Browman & Goldstein, 1993) have established that what sounds like insertion of a segment can be a side effect of gestural timing relations. Based on acoustic evidence from a production experiment with six Turkish speakers, I argue that such gestural timing produces non-lexical vowels in complex onsets in Turkish—previously described as harmonizing epenthetic vowels (Clements & Sezer, 1982, inter alia). Non-lexical vowels occurred in 88.3% of tokens, and usually resembled [ɯ], failing to undergo vowel harmony. Non-lexical vowels are shorter than underlying vowels, and their F1 and F2 values are more affected by the following vowel, suggesting they are more subject to vowel-to-vowel coarticulation. These results support the hypothesis that the inserted acoustic vowels are targetless, created by timing relations between gestures, with the implication that Turkish phonology does not categorically ban complex onsets.

1. Introduction

Cross-linguistically, underlying consonant clusters, particularly in loanwords, often surface with a vowel sound separating the two consonants. In its most familiar form, this vowel sound corresponds to an *epenthetic* vowel—an additional vocalic segment inserted by phonology to repair a consonant cluster that is prohibited by the language's phonology. Like underlying vowels, epenthetic vowels are phonological objects, so they typically do not differ acoustically from underlyingly present vowels, and they generally participate in other phonological processes. In particular, epenthetic vowels crucially participate in syllabification, since they repair illegal syllable structures (Hall, 2011). The position and quality of epenthetic vowels are affected by the range of repairs and structures that are available elsewhere in the language, as well as perceptual factors (Fleischhacker, 2005; Broselow, 2015).

To provide an example, epenthesis of a high vowel occurs in Turkish to repair illegal consonant clusters in codas. These illegal clusters have rising or flat sonority, and occur in Arabic loanwords (Clements & Sezer, 1982). The inserted high vowel (in italics in [1]) appears in the bare form of the word, or when the root is followed by a consonant-initial suffix. It is absent when the consonant cluster is followed by a vowel-initial suffix, such as in the accusative case.

(1)   Coda-repair in Turkish (consonant cluster is bolded, inserted vowels are in italics)

	Root	Nominative	Accusative	Gloss
a.	/sa**br**/	[sɑ.'**b***ɯ***r**]	[sɑ**b**.'**r**ɯ]	'patience'
b.	/dʒe**br**/	[dʒe.'**b***i***r**]	[dʒe**b**.'**r**i]	'algebra'
c.	/bu**rn**/	[bu.'**r***u***n**]	[bu**r**.'**n**u]	'nose'
d.	/ø**mr**/	[ø.'**m***y***r**]	[ø**m**.'**r**y]	'life'

The inserted vowel in illegal coda clusters forms the nucleus of a syllable and allows the final consonant to be syllabified as a simple coda. Turkish stress placement is generally word-final, and like underlying vowels, inserted vowels in coda clusters receive stress when they occur in the final syllable. Coda-repairing vowels are also subject to vowel harmony. The Turkish vowel inventory contains eight phonemes distinguished by [±high], [±back], and [±round]; all three features are relevant to harmony, which affects most suffix vowels. The backness of harmonizing vowels in Turkish is determined by rightward spreading from the nearest vowel in the root. The nearest root vowel also determines the roundness of high harmonizing vowels. Low vowels may trigger rounding harmony but are not targets for it. This harmony process can be seen in the variable realization of the accusative suffix in (1): [ɯ] following /a/, [i] following /e/, and [y] following /ø/. Like the accusative suffix, the inserted vowels in the nominative forms take their backness and rounding from the adjacent root vowel, which indicates that they are targets for vowel harmony. Since the Turkish coda-repairing vowel participates in syllabification, stress-assignment, and vowel harmony, it has to be a phonological object. Thus, it must be epenthetic—a segment inserted during phonology, and mapped to a gesture during articulation.

However, an added vowel sound at the surface does not always correspond to an inserted phonological segment with an accompanying gesture. Studies in Articulatory Phonology (Browman & Goldstein, 1993) have established that what sounds like insertion or deletion of a segment can actually be a side effect of gestural timing relations. When the gestures for adjacent consonants do not overlap, a vowel-like interconsonantal interval can result. The resulting *intrusive vowels* (term adopted from Hall, 2003) can be schwa-like or can 'copy' the quality of an adjacent vowel whose gesture overlaps the interval between consonants.

Intrusive vowels contrast with lexical and epenthetic vowels phonologically, gesturally, and acoustically. Phonologically, intrusive vowels have no corresponding segment, so they cannot participate in phonological processes that target segments, such as vowel harmony and syllabification. Because intrusion does not result in a new vocalic segment, it does not alter syllable structure, and cannot be taken as repair of illegal syllable structures. When vowel intrusion occurs between two consonants, the two consonant gestures are still coordinated with each other. For instance, if vowel epenthesis occurs in complex codas, the maximal syllable for the language might be CVC. But if vowel *intrusion* occurs, then the two final consonants are still part of the same syllable, and in some sense CVCC syllables are permitted. Finally, since intrusive vowels do not form syllable nuclei, they also cannot be targets for stress assignment.

Hall (2003) uses phonological evidence to argue for the intrusive status of inserted 'copy vowels' that repair sonorant-obstruent codas in Scottish Gaelic—they are invisible to syllable-counting in poetry, for example. Similar arguments suggest a gestural-timing origin for Dorsey's Law vowels in Hocank (Winnebago) (Steriade, 1990; Hall, 2003), as well as inserted vowels in Finnish, Dutch, Kekchi, and Mono (Hall, 2003), Spanish (Bradley, 2004; Schmeiser, 2009), and Moroccan Arabic (Gafos, 2002).

Gesturally, intrusive vowels lack corresponding gestures and targets, so they differ articulatorily from phonologically present vowels. **Figures 1** and **2** schematize the gestural sequences in an underlying /CrV/ word that is pronounced as [CVrV]. When epenthesis inserts a segment in phonology, this inserted segment is mapped to an additional gesture during articulation—represented in **Figure 1** by a bolded dashed line.

In contrast, when intrusion occurs, no segment is inserted in phonology, and no gesture is added in articulation, but the relative timing of the /C/ and /r/ gestures produces the percept of an intervening vowel <v> (**Figure 2**). This <v> can sound schwa-like when /V/ overlaps less with the interconsonantal interval (**Figure 2a**), or sound like a copy of the following /V/ when the /V/ gesture overlaps more (**Figure 2b**). Intermediate or more extreme alignments are also possible.

Experiments have exploited these articulatory differences to distinguish intrusive and epenthetic inserted vowels. For example, the intrusive schwas that break up illegal onset clusters like /zg/ in English are gesturally closer to /sk/ than to /sǝk/ (Davidson & Stone, 2003), indicating that the acoustic schwa lacks its own gesture. In contrast, inserted schwas in Dutch, argued by Hall (2003) to be intrusive, have gestural consequences similar to lexical schwa, suggesting that they are epenthetic instead (Warner, Jongman, Cutler, & Mücke, 2002).

Acoustically, since intrusive vowels have no durational target, they are typically shorter than lexical vowels. In addition, since intrusive vowels have no gestural target, their formant values are more affected by coarticulation. Hall and Sue (2018) show that the 'copy-vowels' in Hocank are indeed shorter than lexical vowels. Davidson (2006) shows that intrusive schwas in English speakers' productions of non-native consonant sequences are likewise shorter than lexical schwas, as well as more affected by coarticulation with the following vowel. Vowel intrusion, then, incompletely neutralizes the contrast between /CC/ and /CVC/.

Cross-linguistically, intrusive vowels typically occur across sonorants, share the quality of the vowel that is adjacent across the sonorant, do not contribute a syllable, and are sensitive to speech rate (Hall, 2003; see also Fleischhacker, 2005). These properties characterize complex onset repair in Turkish, in which an underlying consonant cluster optionally surfaces with an acoustic vowel breaking it up, as in (2). The hypothesized intrusive vowels are transcribed between <angle brackets>.
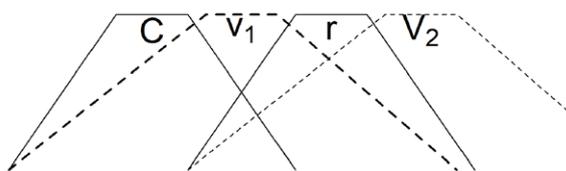
Figure 1: Gestural score for epenthesis.
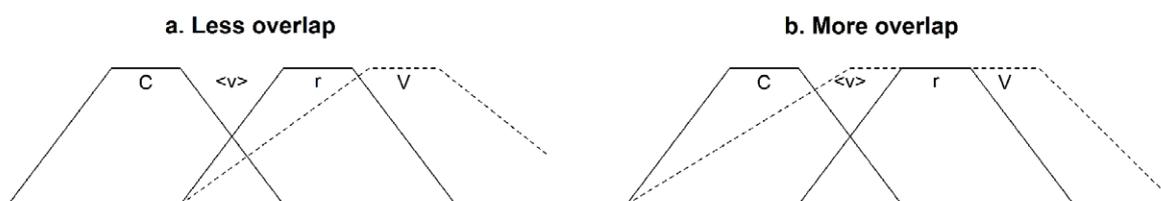
Figure 2: Gestural scores for vowel intrusion.

(2)   Onset-repairing vowel insertion in Turkish (from Clements & Sezer, 1982)

    a. /**pr**ens/ [p<i>rens]   'prince'
    b. /**pr**ova/ [p<u>rova]   'test'
    c. /**br**anda/ [b<ɯ>randa]   'canvas'
    d. /**bl**uʒin/ [b<u>luʒin]~[b<y>lyʒin] 'blue jeans'

The onset-repairing vowels in these examples are invited by a stop + liquid cluster, and their quality is affected by the vowels that are adjacent over the liquid. Onset-repairing vowels can be absent in careful speech (Clements & Sezer, 1982), and are rarely written down. These characteristics suggest that the onset-repairing vowel may be intrusive, an acoustic consequence of the open transition between consonant gestures.

However, previous treatments of Turkish complex onset repair characterize it as the mirror image of complex coda repair. Both non-lexical vowels are described as epenthetic and harmonizing with the neighboring vowel (Yavaş, 1980; Clements & Sezer, 1982; Kaun, 1999; Yıldız, 2010). But where the coda-repairing vowel is obligatorily present in careful speech, casual speech, and in writing, the onset-repairing vowel is only optionally present in both speech and writing. Moreover, where the coda-repairing vowel participates obligatorily in vowel harmony, the onset-repairing vowel reportedly participates in a variable, consonant-dependent fashion (Clements & Sezer, 1982). These differences are explained if onset repair is vowel intrusion, while coda repair is epenthesis.

This paper presents an acoustic production experiment on Turkish onset cluster repair. The results support the hypothesis that vowels in Turkish onset clusters are intrusive, not epenthetic. The duration of the interconsonantal interval (ICI) in Turkish onset clusters is found to have a unimodal distribution, suggesting that the acoustic insertion is a gradient phenomenon, not an optional, categorical process. Moreover, acoustic non-lexical vowels are found to be shorter and more affected by co-articulation with the following vowel than their underlying counterparts. Finally, the formant values of the acoustic inserted vowels in this experiment were generally similar to those of [ɯ], even when vowel harmony would demand [i] or [u]. These results support an interpretation of Turkish onset repair as a gradient gestural phenomenon, in which the release of the initial consonant in the cluster contributes a schwa-like acoustic vowel.

This study contributes in three areas. First, it provides new, controlled Turkish data, by collecting repeated productions by multiple speakers of methodically chosen near minimal sets of words. Second, it probes the phonological status of the Turkish onset-repairing vowel, thereby testing the validity of phonological arguments that have been made on the basis of its behavior. Onset cluster repair is significant for our understanding of both syllable structure and vowel harmony in Turkish. If onset repair is not phonological, then traditional characterization of Turkish syllable structure as maximally CVC(C) needs to be revised, at least for loanwords. In addition, onset cluster repair provides the only counter-evidence to the traditional claim that harmony in Turkish is strictly left to right. If onset repair actually occurs outside of categorical phonology, then it is not actually relevant to harmony. Finally, this study expands the knowledge-base for vowel intrusion by supplying phonetic detail about intrusive vowels that are unusual because they occur in onset clusters (rather than coda clusters), and in a language with vowel harmony.

The remainder of the paper is organized as follows: Section 2 summarizes previous work on onset cluster repair in Turkish. The design of an acoustic production study is presented in Section 3, and its results in Section 4. Section 5 discusses and concludes.

2. Background

2.1. Previous descriptions of Turkish onset cluster repair

In Clements and Sezer's (1982) feature-spreading treatment of harmony and disharmony in Turkish, complex onsets are reported to surface faithfully in careful speech, but be broken up by epenthesis in casual speech. Likewise, Yıldız (2010) describes epenthesis in onset clusters as being in free variation with faithful productions. Both Clements and Sezer (1982) and Yıldız (2010) characterize the onset-repairing vowel as a high vowel whose backness and rounding are determined by regressive harmony with the following vowel. Clements and Sezer (1982) also report that the quality of the onset-repairing vowel varies. Furthermore, unlike the coda-repairing epenthetic vowel, the onset-repairing vowel is reported to always be [+back] after dorsal consonants /k/ and /g/, and to optionally be [–back] following /s/, even in the absence of a [–back] lexical vowel to trigger harmony (Clements & Sezer, 1982). This characterization of onset cluster repair as optional epenthesis (Clements & Sezer, 1982; Yıldız, 2010) predicts that some onset clusters are repaired with phonologically present vowels, with durational and gestural targets like lexical vowels, while other onset cluster tokens contain no vowels, meaning the /CC/ cluster will have a categorically different durational target. In contrast, if onset cluster repair is actually gestural vowel intrusion, the variability is predicted to be gradient. Inter-speaker variation is also predicted: Speakers who are familiar with coordinating the gestures for onset clusters in other languages (e.g., French or English) might be better able to closely coordinate the gestures in Turkish as well.

  Experimental data on Turkish onset cluster repair comes from Yavaş (1980), Kaun (1999), and Bokhari, Durmaz, and Washington (2016). In a reading task (Yavaş, 1980), nonce words that began with consonant-consonant sequences were consistently produced with inserted vowels whose backness reflected the features of the following lexical vowel. Only two of the target words began with obstruent+sonorant clusters, and these two were the source of the only inter-speaker variability in the study. In addition, high round vowels triggered rounding of the inserted vowel, but /o/ did not. Following up on this result, Kaun (1999) presented nine subjects with a list of 109 loanwords beginning with consonant clusters, and asked them what vowel they would say them with. All inserted vowels were high vowels that matched the backness of the following lexical vowel. When the following lexical vowel was [+high, +round], inserted vowels were also consistently round.[1] However, rounding varied between and among speakers when the trigger was low, which was interpreted as a height-agreement effect (Kaun, 1999). Finally, Bokhari et al. (2016) provide the only acoustic study of vowel insertion in Turkish to date. Their production study with four speakers found that coda-repairing vowels did not differ significantly from underlying vowels, while onset-repairing vowels (coded as <i>) had a shorter duration, lower F2, and sometimes a higher F1 than underlying /i/.

2.2. Onset repair in a corpus

To supplement these studies, I conducted a corpus study in the Turkish Electronic Living Lexicon (TELL; Inkelas, Küntay, Sprouse & Orghun, 2000). TELL consists of phonemic transcriptions of 17,500 Turkish lexemes produced by two native speakers of Istanbul Turkish. The data was collected by having these two speakers read through a dictionary and a list of place names, producing each lexeme in a variety of morphological contexts.

---

[1] This contrasts with the standard rounding harmony in Turkish, which is triggered by both low and high vowels, but only targets high vowels.

Of the 415 tokens of word-initial onset clusters in TELL, 70% are transcribed with an inserted vowel. Looking specifically at stop + rhotic clusters, which will be the focus of the production experiment described below, Speaker 1 has 189 input /Cr/ words, of which 135 are transcribed with vowel insertion (71.4% transcribed insertion rate among /Cr/-initial words). The Turkish rhotic, transcribed as /r/ in this paper, is typically realized as an alveolar tap (Göksel & Kerslake, 2005: 9; Lewis, 1967: 7).

Across all cluster types, by far the most commonly transcribed non-lexical vowel in TELL is [ɯ]. Non-lexical front vowels were transcribed in only 31% of the cases where a lexical front vowel trigger was available, contrary to the 100% application of backness harmony reported in Kaun (1999), Yavaş (1980), and Yıldız (2010). In accordance with Clements and Sezer's (1982) claim that front vowels are not inserted after velar consonants, non-lexical front vowels were only transcribed after /k/ or /g/ in a single token (*klişe* 'cliche' [kiliʃe]). In addition, round non-lexical vowels were transcribed in only 36% of the tokens where a [+ round] trigger was present. In line with the results from Kaun (1999) and Yavaş (1980), low round lexical vowels in TELL did not generally trigger rounding in the transcribed inserted vowels, although there were two exceptions where [u] was transcribed as being inserted before /o/. The low rate of transcribed harmony in TELL is surprising in light of some of the descriptions outlined above (the exception being Clements & Sezer, 1982, who predict no insertion at all in careful speech), but fits well with the picture of onset cluster repair as vowel intrusion.

2.3. Discussion of previous work and corpus study

Although prior work largely describes onset cluster repair in Turkish as epenthesis accompanied by vowel harmony, it also reveals differences between onset cluster repair and other epenthesis and harmony in Turkish. Vowel insertion in onset clusters is variable (Clements & Sezer, 1982; Yıldız, 2010; TELL results here), and the inserted vowels may differ acoustically (Bokhari et al., 2016). The harmony that affects onset-repairing vowels is not just the mirror image of the harmony that operates elsewhere in Turkish, because normal rounding harmony in Turkish is not affected by the height or backness of the harmony trigger. In Kaun's (1999) interpretation, the harmonic behavior of onset-repairing vowels reflects a different harmony process, driven by normally inactive constraints from Universal Grammar. The failure of low vowels to trigger harmony is ascribed to a requirement that the trigger and target agree in height. Nonetheless, these results are surprising from an epenthetic perspective, because low vowels are better harmony triggers than high vowels cross-linguistically (Kaun, 1995). This is ascribed to the fact that the perceptual cues to the roundness of low vowels are weaker than the perceptual cues to the rounding of high vowels—grossly speaking, high vowels are rounder than low vowels. This articulatory fact suggests that coarticulation with a high round vowel is more likely to produce an impression of rounding on an intrusive vowel than coarticulation with a low round vowel—which is exactly the pattern of (apparent) harmony that emerges for the Turkish onset-repairing vowel (Kaun, 1999; Yavaş, 1980; corpus results above).

One possible problem with interpreting onset cluster repair in Turkish as vowel intrusion is that vowel insertion is reported even in clusters containing two obstruents, particularly in Yavaş (1980). According to Hall (2003, 2006), vowel intrusion occurs only across sonorant consonants, whose gestures are better able to overlap with vocalic gestures. However, other work does report gesturally-driven vowel intrusion in obstruent + obstruent clusters (e.g., Gafos, 2002; Davidson & Stone, 2003; Davidson, 2006; Davidson, 2010). While it is possible that intrusion is only occurring across sonorants in Turkish onsets, and optional epenthesis is repairing obstruent + obstruent clusters, it also seems likely

that intrusion across obstruents is possible, contra Hall's (2003) criteria, and it simply is less common, whether for the articulatory reasons Hall (2003) points to, or because a obstruent + sonorant cluster is less perceptually altered by vowel intrusion than an obstruent + obstruent cluster is (Fleischhacker, 2005).

3. Production experiment

To address the lack of data on acoustic detail, intraspeaker variation, and the effect of the surrounding context on onset cluster repair in Turkish, I conducted a production study. The experiment is designed to: (1) establish whether apparent insertion in Turkish is a gradient or a categorical process, by examining the duration of the interval between C and /r/; (2) determine the rate of acoustic insertion in onset clusters and the degree to which frontness or rounding spreads to the inserted vowel; (3) look for acoustic differences between lexical and non-lexical vowels. The experiment had a 2 by 3 by 3 by 2 design. The primary factor manipulated was the underlying syllable structure of the target word: beginning with a stop + /r/ onset cluster (/Cr/), or beginning with a simple onset followed by an underlying vowel and /r/ (/Cvr/). The /Cvr/ words were included as controls so that non-lexical vowels in /Cr/ words could be compared to lexical vowels. Although vowel insertion is also reported to occur in other clusters (including /s/ + stop, obstruent + /l/), /Cr/ clusters were chosen for the experiment because insertion is transcribed at a higher rate in /Cr/ clusters (71% in TELL) than in /sC/ clusters (42% in TELL). In addition, surface harmonic effects resulting from vowel overlap are more likely to occur across a sonorant like Turkish /r/ (phonetically a tap) than across a stop (Hall, 2003, 2006).

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/— voiced stops were chosen to avoid aspiration) and three vowels (/i/, /a/, /o/)[2] were included. Finally, both real, familiar words and completely unfamiliar nonce words were included, to check that insertion is a fully productive process.

3.1. Materials

A list of real and nonce words beginning with stop + /r/ clusters was constructed (**Table 1**, Experimental columns). Target words take the form $/C_1(v_1)rV_2C_2.../$. Within each C_1-V_2 condition, C_2 was matched for major place of articulation. Stress was also controlled so that syllables that would be compared were all unstressed; stress falls on the final syllable (V_2 or later) in all words. Finally, the number of syllables was also controlled, such that all output forms in a C_1-V_2 condition are predicted to have the same number of syllables (not counting potential inserted vowels as syllabic).

Real /Cr/ words were chosen to be familiar, where possible. Familiarity was determined on the basis of a familiarity-rating survey conducted with three native speakers of Turkish (1 female, 2 male; ages 28–63), who did not participate in the experiment otherwise. Participants were asked to rate the familiarity of the words on a five-point scale, where 1 meant "I don't know this word at all" and 5 meant "I use this word regularly or learned it as a young child." Instructions were presented in Turkish. A word was considered familiar if it received an average rating of at least 4 on the survey, with no participant giving it a rating of 1 or 2. Unfortunately, in the /dri-/ and /gro-/ conditions, no sufficiently familiar Turkish word was found, so the highest-rated available word was selected even though

---

[2] Originally /u/ was included as the third V_2 value, rather than /o/, but no sufficiently familiar words of the shape /bru-/ could be found, and so /o/ was selected instead. In some theories, /o/ is considered to be a better trigger of rounding harmony than /u/ (Kaun, 1995), so /o/ also provides a better test of whether the quality of the inserted vowel is determined by phonological vowel harmony or by phonetic coarticulation. Additionally, /o/ as a non-high vowel provides more information about whether acoustic inserted vowels seem to share the height of the lexically present vowel, or only its backness and rounding.

Table 1: Stimuli. Unglossed items are nonce words. An asterisk following a word indicates that it is also being used as a /CVr/ match for a /Cr/ word in the real word condition, since no appropriately shaped real word could be found. Familiarity ratings for real /Cr/ words are shown in parentheses.

C_1	V_2	Experimental		Control	
		Real /Cr/ *word (familiarity) 'gloss'*	Nonce /Cr/	v_1 = <ɯ>	v_1 ≠ <ɯ>
b	/i/	bri.fing (4) 'briefing'	bri.mi.ti	bɯ.ri.pis	bi.ri.m-in 'unit.your' bi.ri.bis
	/a/	bran.ʃ-ɯ (4.33) 'subject.ACC'	brat.tʃi.ten	bɯ.ran.dʒɯ*	–
	/o/	bro.ʃyr (4.67) 'brochure'	bro.ʒør.le	bɯ.ro.ʒyn*	bu.ro.tʃyp*
d	/i/	drip.ling (1) 'dribbling'	drip.li.ke	dɯ.rib.le*	di.rim.-ler 'life.PL' di.rib.rit
	/a/	dra.ma (4) 'drama' *or* dra.m-a (3.7) 'drama.DAT'	dra.fa	dɯ.rap*	–
	/o/[3]	bor.dro-m (4) 'payroll.my'	lor.dro.pur	gar.dɯ.rop 'wardrobe'	nor.du.rof*
g	/i/	grip (5) 'influenza'	gri.vi	gɯ.rif*	gi.rim 'penetration' gi.riv
	/a/	gram (5) 'gram'	gra.bɯ	gɯ.rap*	–
	/o/	gro.s-u (2.67) 'gross.ACC'	gro.dol	gɯ.ron*	gu.rot*

ratings were quite low (1 for *dripling* 'dribbling' [as in basketball]; 2.67 for *gros* 'gross' [as opposed to net]).

Control words of the form /CVrV/ were created for every condition (**Table 1**, Control columns) so that non-lexical vowels in /Cr/ words could be compared to lexical vowels in the same context. The number of relevant /CVrV/ control words varies depending on the identity of V_2. The insertion of [ɯ] is attested before all qualities of V_2, so control /CVrV/ words with v_1 = /ɯ/ were created in every V_2 condition (i.e., /Cɯri/, /Cɯra/, and /Cɯro/). In addition, [i] is reported to be inserted before /i/, and [u] is reported to be inserted before /o/, so I created control words where v_1 and V_2 were both [–back] or [+round] (/Ciri/ and /Curo/). Only [ɯ] is reported before /a/, since it is already harmonic for backness and rounding, so there are fewer relevant control words in /a/ conditions.

While [ɯ-a] sequences are harmonic for both backness and rounding, [ɯ-i] sequences are disharmonic for backness, and [ɯ-o] sequences, for rounding. These disharmonic sequences are unattested as underlying sequences (except for *gardurop* 'wardrobe') in the corpora and dictionaries I consulted. This gap in the lexicon suggests that Turkish phonology prohibits these particular disharmonic vowel sequences. Onset-repairing vowel insertion creates them in surface forms, however. Therefore, the necessary /Cɯri/ and /Cɯro/ controls had to be nonce words, and it was not possible to maintain distinct real and nonce conditions

---

[3] A note about the /dro-/ cell: In all other C-V conditions, the consonant-cluster of interest is word-initial. But in the dro- condition, the cluster appears word-internally (/bordrom/ 'payroll.my'). This word was selected in order to maintain the same environment for the cluster as for the underlyingly present vowel, in order to be able to use the real word gardɯrop 'wardrobe" for the /Cɯro/ control word. Ultimately, however, this turned out to be a mistake, because the /rdr/ sequence that was intended as a coda /r/ followed by a complex onset was instead interpreted as a complex coda followed by a simplex onset. Consequently, the /dro/ condition was omitted from the analysis.

in the control words. Instead, nonce control words were included in all conditions, and real words were also included when they existed, resulting in different numbers of control words depending on the condition. No familiarity ratings for control /Cvr/ words were obtained, since so many nonce /Cvr/ words were included, and since I expect familiarity levels to have no significant impact on the articulation of a lexically present v_1.

Stimuli are shown in **Table 1**. There are 24 words each in the real and nonce word conditions, but 11 nonce /Cvr/ words overlap between the two conditions, so the total number of distinct target words is 37.

In addition, 17 fillers (**Table 2**) were included, for a total of 54 target words. Because so many experimental items are nonce words, primarily real words were selected as fillers— mostly borrowings from English or French since all the familiar real words are borrowings.

Both target and filler words were presented in the carrier sentence in (3), which includes slots for two target words. The sentence was designed to elicit contrastive focus on the target words, to further enhance the carefulness of the elicited speech.

(3)　　Bana　　X deme,　　bana　　Y de.
　　　　me.DAT　X say.NEG,　me.DAT　Y say.
　　　　"Don't say X to me, say Y to me."

Since the structure of the carrier sentence elicits an expectation of structural parallelism (that X and Y will be of the same grammatical category and case), X~Y pairs with the same case were selected. Also, within a given sentence, X and Y were either both nonce or both real. To control for the possibility that prosodic factors would create a confounding difference in articulation between X and Y, half the repetitions employed an X-Y order, and the other half employed a Y-X order.

3.2. Participants

Six native speakers of Turkish (3 female: S4, S5, S7) were recruited from the University of California at Santa Cruz community. (A seventh [S1] participated in the pilot experiment, after which the design was significantly revised, so her data are not discussed.) S3 is bilingual in French and Turkish, so language effects may complicate the interpretation of his data. S6 lived in New Jersey, USA, for a year (age 4–5), but in Turkey otherwise. 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. Participants were paid $20 for their time.

3.3. 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 pronounce words. Recordings

Table 2: Fillers.

C$_1$	V = /e/	V2 = /u, o/	V = /a/
Labial	merimit (nonce)	provizjon 'commission'	marɯp (nonce)
	meteoroloʒi 'meteorology'		paralelkenar 'parallelogram'
Coronal	negatif 'negative'	tuvalet 'toilet'	tablo 'painting'
	neptyn 'Neptune'	turnike 'turnstile'	tansijon 'blood pressure'
Dorsal	kervan 'caravan'	kuafør 'hair dresser'	kakao 'cocoa'
	geometri 'geometry'		karton 'carton'

were made in a sound-attentuated booth using a shotgun microphone with a USB pre-amplifier. Subjects were asked to practice reading the instructions to get comfortable speaking with 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 subjects 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 sentences were randomized together, without any blocking of real vs. nonce words. 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 had not identified the research question being investigated.

Acoustic annotation of the v_1 interval was conducted in Praat (Boersma & Weenink, 2015) using TextGrids. The left edge of the interconsonantal interval (ICI) was marked from the beginning of the C_1 release burst, identified by a dramatic increase in amplitude. The right edge of the ICI was identified by the decrease in amplitude accompanying the onset of /r/. The ICI was further subdivided into the burst + VOT (annotated as 'burst') and v_1, where v_1 was identified as the portion of the ICI that had high amplitude periodicity and formant structure. Sometimes no such formant structure occurred. Less commonly, high amplitude periodicity with formant structure sometimes occurred throughout the ICI. Representative spectrograms are shown in **Figure 3** (underlying vowel), **Figure 4** (non-lexical vowel), and **Figure 5** (cluster with no vowel).

Measurements of F1 and F2 were taken at the midpoint of v_1 using a Praat script, and converted from Hertz to Bark using the formula from Traunmüller (1990). I excluded nine tokens which stood out as outliers when the vowels were plotted; their F1 in Bark was less than 1.9 (*buroʒyn, burotʃyp,* and *gurot* from S3; *branʃu* from S7) or greater than 5.2 (*bratʃiten, broʒørle,* and *driplike* from S3; *dirimler,* S4; *buroʒyn,* S7). Five tokens in which V$_2$ was mispronounced were also excluded (three repetitions of *bratʃiten* from S4 and one each of *driplike* from S4 and *guron* from S3, all nonce words). Finally, the *dro-* condition was excluded because the /dr/ cluster occurred word-medially and was not syllabified as a complex onset in many cases. All other clusters were word-initial. The resulting dataset contained 936 tokens from six speakers.
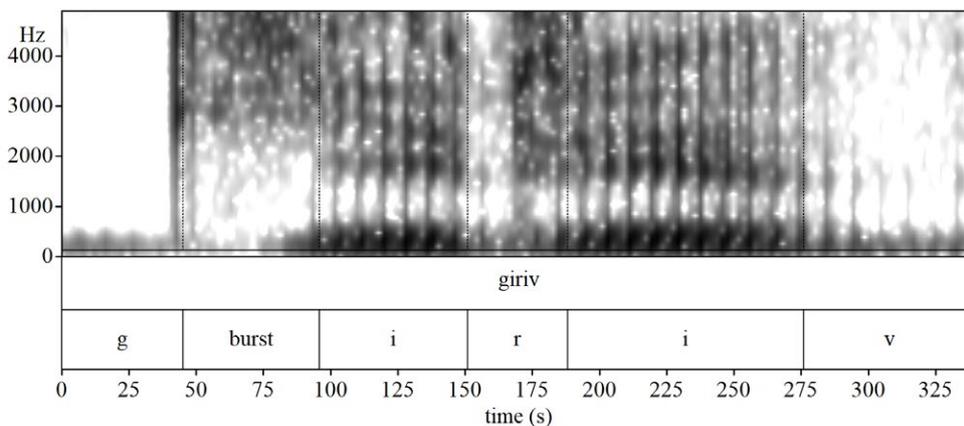
Figure 3: /Cvr/ token with an underlying vowel (from S3).
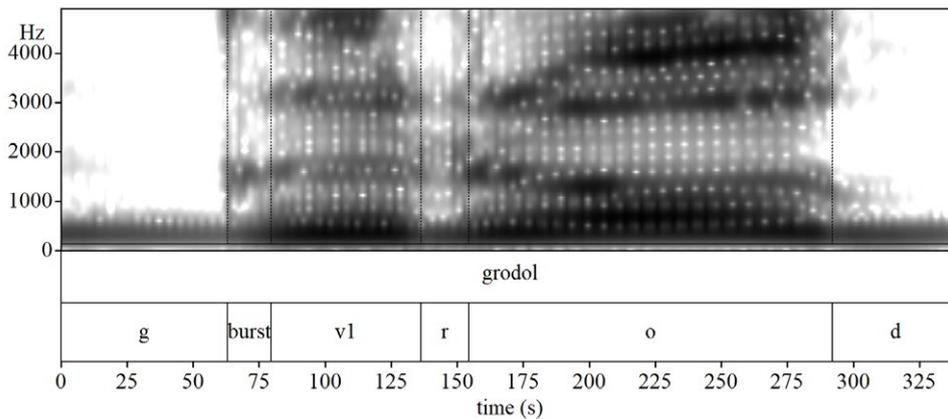
Figure 4: /Cr/ token containing an acoustic non-lexical vowel (from S4).
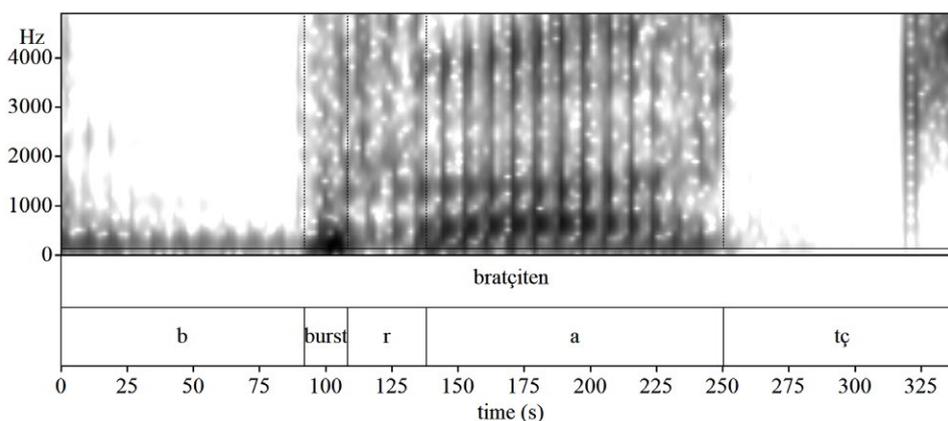
Figure 5: /Cr/ token with no acoustic insertion (from S3).

4. Results

No differences were found between real/familiar words and nonce words, so real and nonce words are treated together throughout the analysis. Acoustic analysis reported in this section finds that onset cluster repair is variable and gradient (Section 4.1). Non-lexical vowels tend to be acoustically [ɯ]-like, rather than conforming to vowel harmony (Section 4.2). However, non-lexical vowels display significant differences in duration, F1, and F2 from both harmonic lexical vowels (Sections 4.3–4.4) and from lexical /ɯ/ (Section 4.5).

4.1. Gradience in onset repair

Vowel intrusion results from gradient gestural alignment, so if onset cluster repair is vowel intrusion, ICI durations are predicted to have a unimodal distribution of durations. On the other hand, vowel epenthesis reflects a categorical insertion process, so if onset cluster repair is optional epenthesis, as traditionally described, ICI durations are predicted to have a bimodal distribution (one mode for insertion and one mode for no insertion).

Using R (R Core Team, 2016), the distributions of two duration-measures for underlying clusters were plotted (**Figure 6**): the ICI, and the portion of the ICI with high amplitude periodicity with formant structure (which I refer to as the vowel). The distribution of lexical vowels is also plotted for comparison (dashed line). Both distributions appear unimodal, suggesting that acoustic insertion is a gradient process, not an optional categorical one. (The secondary mode in the smoothed density curve for vowel durations is an artifact

of coding the underlying clusters that were produced with no vowel as having a vowel duration of 0 ms; of course no negative durations are possible.) The ICI is shorter in underlying clusters than in lexical vowels, although these are quite short as well (mean ICI duration = 74.1 ms, mean vowel duration = 53.6 ms).

For purposes of comparing acoustic non-lexical vowels to lexical vowels, a vowel-duration threshold (shown in red in **Figure 6**) was established at 20 ms,[4] since the histogram of non-lexical vowels reveals a sharp change between the 20–30 ms and 10–20 ms bins, and all but two underlying vowels are longer than 20 ms. Clusters produced with at least 20 ms of high amplitude periodicity with formant structure were coded as containing an acoustic non-lexical vowel. With this criterion, acoustic vowel insertion occurs in 88.3% of the underlying clusters, with the insertion rate varying between subjects (**Figure 7**).

4.2. Vowel plots and F1~F2 of v_1

Using R (R Core Team, 2016) and the 'ellipse' package (Murdoch & Chow, 2018), both lexical and non-lexical v_1 tokens were plotted in F1~F2 space (**Figure 8**).

Consistent with Kiliç and Öğüt's (2004) report that /ɯ/ is more mid/central than other Turkish back vowels, /ɯ/'s F1 and F2 values are intermediate between those for /i/ and for /u/. Most non-lexical vowels (represented by open circles) lie within the distribution of lexical /ɯ/, and few of them lie within the distribution of lexical /i/. Acoustically speaking, then, harmony does not seem to have applied to the onset-repairing vowels in a categorical, consistent way, contra previous descriptions. These acoustic differences between lexical and non-lexical vowels are investigated below.
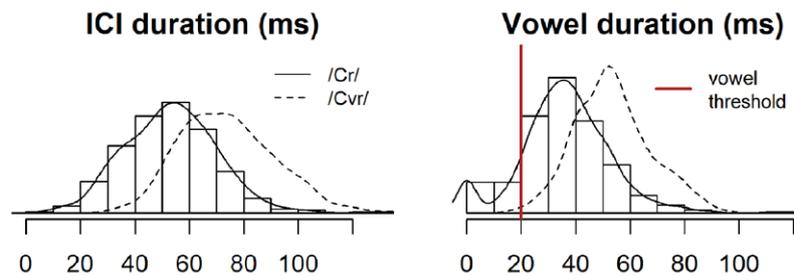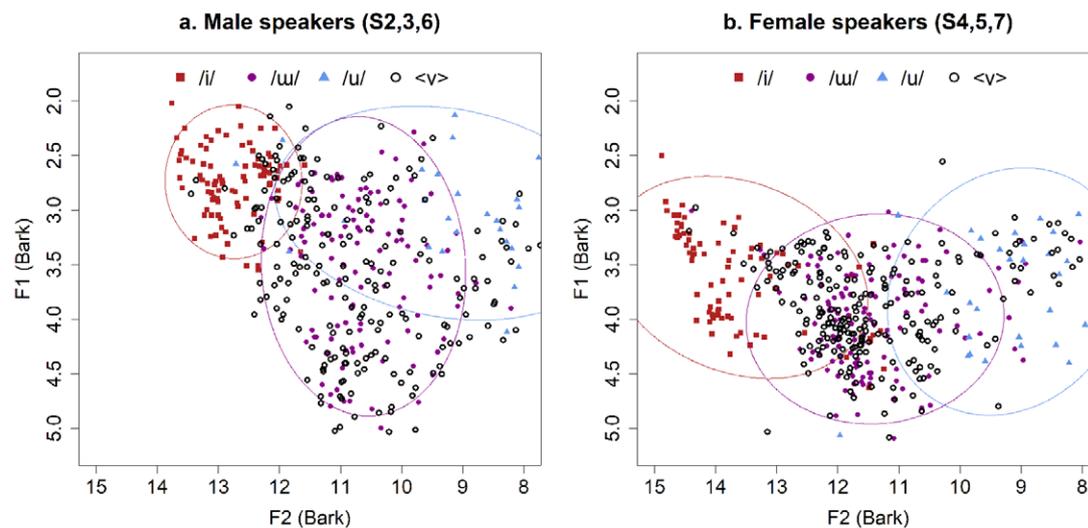

Figure 6: Duration of ICI and of v_1.
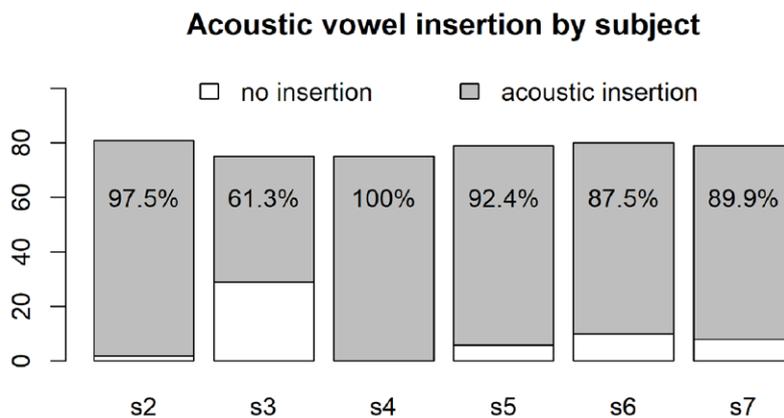
Figure 7: Acoustic insertion in onset clusters by subject.

---

[4] If we take a more conservative approach and place the threshold midway between the mean of the lexical vowel duration distribution (57.4 ms) and the mean of a hypothesized no-insertion distribution centered on 0 (i.e., at 28.6 ms), all results are essentially the same.

Figure 8: Vowel plots. Non-lexical vowels are plotted as open circles.

4.3. Acoustic differences between lexical and non-lexical vowels

If the non-lexical vowels are intrusive, as hypothesized here, they will lack the durational and gestural targets associated with true vowels, and are predicted to be shorter and more subject to coarticulation than lexical vowels. According to the standard epenthetic theory of onset-cluster repair in Turkish, non-lexical vowels are subject to backness and rounding harmony. For purposes of testing that hypothesis, I treat the non-lexical vowels accordingly: as <i> before /i/, <ɯ> before /a/, and <u> before /o/ (cf. Clements & Sezer, 1982; Yıldız, 2010).

Linear mixed effects models of duration, F1, and F2 were computed using R (R Core Team, 2016) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2016). All models included fixed effects for the preceding consonant, the category of the following vowel (V_2), and the hypothesized category of v_1, as well as random intercepts for subject and item. Models representing the intrusive hypothesis additionally included the lexical status of v_1 (whether the word underlyingly begins with /Cr/ or /CVr/), along with one or more interactions. Since the epenthetic hypothesis predicts no acoustic differences between lexical and non-lexical vowels, the models representing this hypothesis did not include v_1's lexical status as a factor. Models are shown in the Appendix.
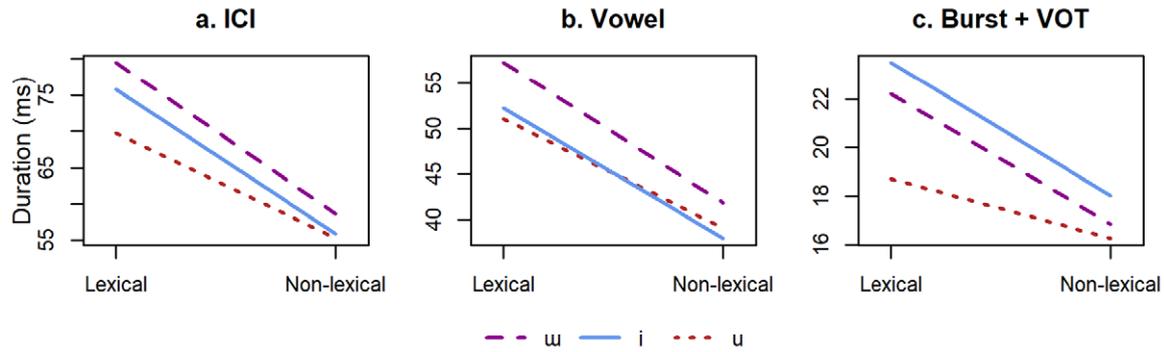
4.3.1. Duration

Three separate measures of duration were analyzed: the duration of the whole interconsonantal interval, the vocalic portion of the ICI, and the burst combined with any additional positive VOT. For all measures, models that included v_1's lexical status performed better than models that did not in maximum likelihood ratio tests (all $ps < 0.001$).

4.3.1.1 Duration of the interconsonantal interval (ICI)

In the best model of ICI duration (`duration.ICI.model`), lexical status had a significant main effect, with the ICI being significantly shorter in non-lexical vowels than lexical vowels ($\beta = -9.01$, $SE = 2.74$, $p < 0.005$). There was also a significant interaction between lexical status and v_1 category—<i> and <u> are shorter than their lexical counterparts by an additional 12.19 ms ($SE = 3.5$, $p < 0.005$) and 10.97 ms ($SE = 4.39$, $p < 0.05$) respectively. The interaction is visualized in **Figure 9a**.

In addition, though less relevant to the hypotheses of this paper, the ICI is longer before /i/ than before /a/ ($\beta = 11.20$, $SE = 3.01$, $p < 0.005$), perhaps reflecting a trade-off effect where a longer V_2 like /a/ results in a shorter v_1. ICI duration is also significantly

Figure 9: Interaction of lexical status and v_1 category for three different measures of duration. Each line is a different v_1 category.

shorter after /b/ than /d/ ($\beta = -4.97$ ms, $SE = 1.9$, $p < 0.05$) and longer after /g/ than /d/ ($\beta = 10.27$ ms, $SE = 1.85$, $p < 0.001$), as predicted by previous work on place effects on VOT (Cho & Ladefoged, 1999) and gestural coordination (Yip, 2013).

4.3.1.2 Duration of v_1

Analysis of v_1, the portion of the ICI that has high amplitude periodicity with formant structure, again found significant main and interaction effects of lexical status (`duration.vowel.model`). In non-lexical vowels, the vocalic interlude is shorter ($\beta = -7.98$, $SE = 2.09$, $p = 0.001$). As illustrated in **Figure 9b**, non-lexical <i> and <u> are particularly short again ($\beta = -8.38$, $SE = 2.70$, $p < 0.01$; $\beta = -7.09$, $SE = 3.37$, $p < 0.05$). Also, like the ICI, the vowel is longer before /i/ than /a/ ($\beta = 9.20$, $SE = 2.31$, $p < 0.001$), as well as slightly longer after /g/ ($\beta = 4.54$ ms, $SE = 1.41$, $p < 0.005$) than after /d/.

4.3.1.3 Duration of burst+VOT

In the analysis of the duration of the consonant burst plus any additional positive VOT (`duration.burst.model`), there was a significant main effect of lexical status ($\beta = -3.89$, $SE = 1.02$, $p < 0.001$). Unsurprisingly, burst durations were shorter for /b/ ($\beta = -4.80$, $SE = 1.25$, $p < 0.001$) and longer for /g/ ($\beta = 5.52$, $SE = 1.24$, $p < 0.005$) than /d/. In addition, burst+VOT was longer in /i/ than /ɯ/ ($\beta = 3.60$, $SE = 1.42$, $p < 0.05$), perhaps due to the tongue's higher position in /i/. The interaction of lexical status and v_1 category did not reach significance, but is shown in **Figure 9c** for comparison to the other duration measures.

4.3.2. F1

The hypothesis that non-lexical vowels are intrusive also predicts differences in their formant values. Formant values were measured at the midpoint of the high amplitude portion of the ICI with periodicity and formant structure. The best model of F1 (`harmony.F1.model1`) included an interaction between lexical status and hypothesized v_1 category, and outperformed the epenthetic model that excluded lexical status ($\chi^2(3) = 25.6$, $p < 0.001$). The interaction between lexical status and hypothesized v_1 category was significant ($\beta = 20.05$, $SE = 9.55$, $p < 0.05$). The higher F1 of non-lexical <i> suggests that it lacks lexical /i/'s [+high] target.

The model shows the expected main effects of surrounding context: F1 is significantly lower for /i/ ($\beta = -56.47$ Hz, $SE = 6.72$, $p < 0.001$) than for /ɯ/, which is known to be lower than /i/ in Turkish (Kiliç & Öğüt, 2004); F1 is also lower after /g/ than /d/ ($\beta = -18.43$, $SE = 5.01$, $p < 0.005$). The effect of the following vowel was also significant: F1 is lower when /i/ follows ($\beta = -18.35$, $SE = 8.17$, $p < 0.05$) and when /o/ follows ($\beta = -38.43$, $SE = 9.63$, $p < 0.001$), compared to /a/.

4.3.3. F2

The best model of F2 (`harmony.F2.model1`) also included an interaction between lexical status and hypothesized v_1 category, and it outperformed the epenthetic model that excludes lexical status in maximum likelihood ratio test ($\chi^2(3) = 32.58$, $p < 0.001$). There was a significant interaction between lexical status and v_1 category, showing that F2 is lower for non-lexical <i> than for underlying /i/ ($\beta = -320.52$ Hz, $SE = 75.33$, $p < 0.001$). This suggests that non-lexical <i> is backer than lexical /i/, perhaps lacking /i/'s [+front] target.

As expected, though less relevant to the hypotheses of this paper, the model also shows that F2 is higher in /i/ than /ɯ/ ($\beta = 393.70$ Hz, $SE = 42.44$, $p < 0.0001$) and lower in /u/ than /ɯ/ ($\beta = -202.93$ Hz, $SE = 76.70$, $p < 0.05$), and that a following /i/ raises F2 ($\beta = 162.94$ Hz, $SE = 62.35$, $p < 0.05$).

4.4. Assuming frontness harmony but no rounding harmony

The models above found significant differences between non-lexical vowels and harmonic lexical vowels in the same context. However, as discussed above, some previous experiments suggest that rounding harmony in onset cluster repair may only be triggered by high vowels (Yavas, 1980; Kaun, 1999). To take this possibility into account, non-lexical vowels were recoded as <i> before /i/ and <ɯ> before /a/ and /o/. Modeling of F1 and F2 under these assumptions recapitulated the effects described above, with non-lexical <i> having a higher F1 (`harmony.F1.model2`: $\beta = 21.03$, $SE = 7.56$, $p < 0.001$) and lower F2 than /i/ (`harmony.F2.model2`: $\beta = -248.63$, $SE = 70.49$, $p < 0.005$), as well as the expected effects of preceding consonant and following vowel.

4.5. Assuming no harmony: Comparing non-lexical vowels to lexical /ɯ/

Logically, the differences between lexical and non-lexical vowels reported above could also result from epenthesis applying but harmony not applying, in which case all epenthetic vowels would be [ɯ]. To address this possibility, non-lexical vowels were again recoded, this time treating all of them as <ɯ>, and the analyses above were repeated. Data was subsetted to exclude /i/ and /u/, since these were no longer relevant.

4.5.1. F1

The best model of F1 assuming no harmony (`noharmony.F1.model`) includes an interaction between lexical status and V_2. This model was significantly better than the epenthetic model that did not include lexical status as a factor ($\chi^2(3) = 16.12$, $p < 0.005$). F1 is lower in non-lexical vowels followed by /i/ ($\beta = -34.04$, $SE = 9.42$, $p < 0.005$). This interaction effect suggests greater anticipatory coarticulation in the non-lexical vowel, since a following /i/ lowers F1 in non-lexical vowels more than in lexical /ɯ/.

Less relevantly, the model also shows main effects of a preceding /g/ ($\beta = -20.49$, $SE = 5.30$, $p < 0.005$), following /i/ ($\beta = -17.95$, $SE = 7.46$, $p < 0.05$), and following /o/ ($\beta = -38.79$, $SE = 8.84$, $p < 0.001$).

4.5.2. F2

The best harmony-free model of F2 (`noharmony.F2.model`) performed better than the epenthetic model in a maximum likelihood test ($\chi^2(3) = 20.16$, $p < 0.001$), and includes an interaction between lexical status and place of the following vowel. Before /o/, non-lexical vowels have a lower F2 than lexical vowels ($\beta = -158.13$, $SE = 49.34$, $p < 0.01$)—further evidence that non-lexical vowels are more affected by anticipatory coarticulation. The model also shows the expected main effects from vowel and consonant context.

4.6. Summary of acoustic differences

Model comparison found that the lexical status of v_1 significantly improved model performance for duration, F1, and F2. Non-lexical vowels are shorter than their underlying counterparts, a result predicted if non-lexical vowels are not true vowels, only the acoustic consequence of an open transition between consonant gestures, which has no durational or acoustic target. In addition, non-lexical vowels are acoustically intermediate between the harmonizing vowels /i/ and /u/ and the non-harmonizing /ɯ/ in their F1 (**Figure 10**) and F2 (**Figure 11**). This means that acoustic differences between lexical and non-lexical vowels are found regardless of whether harmony is assumed to have applied.

The differences between lexical and non-lexical vowels are particularly clear before /i/, where non-lexical vowels had higher F1 and lower F2 than lexical /i/, but lower F1 and higher F2 than lexical /ɯ/ (**Figures 10a** and **11a**). This suggests non-lexical vowels are more centralized than /i/ but also more affected by anticipatory raising and fronting for the following /i/ than lexical /ɯ/. Likewise, non-lexical vowels before /o/ had a significantly lower F2 than lexical /ɯ/ (**Figure 11c**). This suggests that non-lexical vowels are more affected by anticipatory rounding for /o/. Both observations are compatible with the hypothesis that the non-lexical vowels are targetless.

5. Discussion

This paper has presented acoustic evidence that the non-lexical vowels in underlying onset clusters in Turkish result from gradient, gestural intrusion, and consequently lack durational and gestural targets. In the production experiment, 88% of underlying onset clusters in Turkish are produced with an acoustic inserted vowel. Though this high rate of acoustic insertion is contrary to Clements and Sezer's (1982) report that vowel insertion does not occur in careful or formal speech, it is consistent with the overall landscape of data
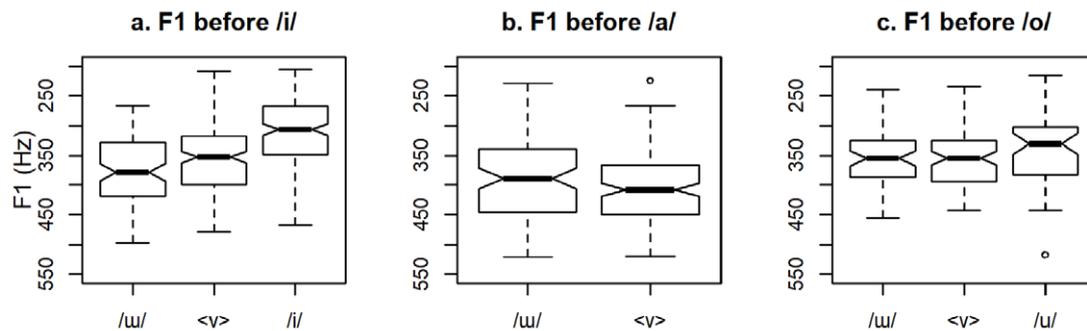
Figure 10: Effect of lexical status on F1. Non-overlapping notches indicate significant differences in group medians.
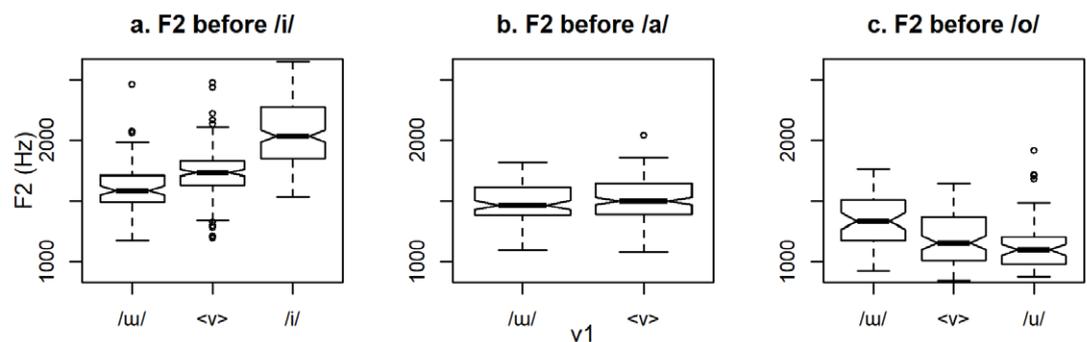
Figure 11: Effect of lexical status on F2. Non-overlapping notches indicate significant differences in group medians.

on Turkish onset cluster repair, since all other work on this topic reports plenty of insertion in laboratory speech, and does not mention an effect of speech style (Yavaş, 1980; Kaun, 1999; Yıldız, 2010; Bokhari et al., 2016). The ICI in words with underlying clusters has a unimodal duration distribution. This is predicted by a gestural account, where the duration of the ICI is determined by the degree of gestural overlap, not by an optional durational target. This indicates that apparent insertion is a gradient process that can 'apply' to a range of degrees, not a categorical but optional process as previously described.

The quality of non-lexical vowels also seems to be gradiently determined by the surrounding gestural context. Non-lexical vowels are acoustically intermediate between harmonizing and non-harmonizing lexical vowels, with F1 and F2 differences being most significant before /i/. These acoustic differences show that the vowels appearing in onset clusters are definitely not participating in backness harmony. This implies they are not participating in rounding harmony, either, although the acoustic differences between lexical /u/ and the non-lexical vowels did not reach significance.

To summarize, vowel intrusion does not completely neutralize the distinction between /CC/ and /CVC/ in Turkish. Rather, the non-lexical vowels in Turkish onset clusters are shorter than lexical vowels; are more affected by the surrounding context; and do not participate in vowel harmony. Moreover, the Turkish lexicon lacks the structures that would be created if the intrusive vowel were taken to be true inserted [ɯ] (i.e., forms containing underlying disharmonic sequences [ɯ i] and [ɯ o]—see Section 3.1), suggesting that the Turkish grammar actually rules out such sequences. These observations argue that non-lexical vowels in Turkish onset clusters are intrusive vowels. Lacking their own gestural targets, the acoustics of these intrusive vowels are determined by their context. There is no insertion of a vowel gesture even in clusters that are produced with an acoustic vowel between the two consonants; instead, this intrusive vowel represents a period when the closure of the first consonant has been released but the closure of the second consonant has not been completed. Meanwhile, the tongue body is already moving toward the following vowel's target, such that the formant values during the ICI are shaped by that target.

This interpretation is readily represented in an Articulatory Phonology (Browman & Goldstein, 1993) framework, where gestures within a syllable are coordinated together in time. This kind of gestural coordination can be represented in the grammar, as in Gafos (2002) and Hall (2003). For example, the gestural coordination that produces vowel intrusion in Turkish onset clusters could be modeled with a constraint aligning the onset of C_2 with the release of C_1. (Two variations on this coordination relation are depicted in **Figure 2**.) This interpretation and the specific gestural alignment(s) could be verified with a gestural study of tongue movements during the production of Turkish onset clusters, which is ongoing (see preliminary results in Bellik, to appear).

Although the gestural coordination that produces intrusive vowels seems to be grammaticized in some languages (Gafos, 2002; Hall, 2003), it is unclear whether this is the case in Turkish, or whether speakers are aiming for a closer transition between consonants in the cluster, but missing that target. Interspeaker variation found in this study (discussed below, Section 5.2) suggests that, at the very least, Turkish speakers vary in the gestural coordination they employ in onset clusters.

5.1. Implications for harmony and syllable structure in Turkish

If onset cluster repairing vowels arise from gestural timing relations, rather than being epenthetic, then their behavior should not be used as the basis for arguments about the segmental phonology of Turkish, particularly vowel harmony. An intrusive vowel cannot be a target for phonological harmony since it is not a phonological object. This suggests that the reasoning behind studies like Kaun (1999)—where the harmonic behavior of

the onset-repairing vowel is used to make claims about speakers' access to phonological constraints that are not active in the native lexicon—must be re-evaluated. Kaun (1999) may bear on the phonetic basis for phonological constraints, rather than the phonology of Turkish vowel harmony per se.

In addition, the non-harmonizing behavior of the inserted vowel cannot be used to bolster the traditional understanding of vowel-harmony in Turkish as a strictly left-to-right process (e.g., Lees, 1966; Underhill, 1986), since an intrusive vowel could never be a target for phonological harmony anyway; neither can its occasional harmonic acoustics—actually due to coarticulation—be attributed to the emergence of a normally invisible right-to-left harmony pattern.

The phonological status of the vowels in Turkish onset clusters is also relevant to our understanding of Turkish syllable structure. If the onset-repairing vowel is not epenthetic, then 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. Future work could test this claim by probing Turkish speakers' mental representations of onset clusters with a syllable-counting task, or by examining text-setting of these non-lexical vowels in music.

5.2. Interspeaker variation and cross-linguistic implications

Given that all onset clusters in Turkish come from loanwords, an anonymous reviewer asks whether it might be the case that the Turkish phonological grammar originally prohibited onset clusters and repaired them with epenthesis, and has changed (or is changing) to permit onset clusters, even if they are realized with an open transition between the consonants. We can also consider the possibility that the transition went in the opposite direction, from initially attempting to produce borrowed onset clusters with a foreign-like gestural coordination, to later producing intrusive vowels, and finally toward reanalyzing the intrusive vowels as epenthetic and integrating the loanwords into the native phonological grammar, which prohibits onset clusters. A full discussion is beyond the scope of this paper, but I would like to propose that both scenarios played out in different segments of the population. It seems likely that there has always been variation in Turkish speakers' realization and representation of onset clusters in loanwords, based on individuals' degree of exposure to the source languages. Post-hoc examination of the inter-speaker variation in this experiment provides tentative support for this: Synchronically, the degree of exposure to languages with onset clusters seemed to predict the degree to which clusters contrasted with /CVC/ sequences. Speakers roughly fall into three groups: categorical differentiators, gradient differentiators, and neutralizers, echoing the pattern in Hall (2013).

First, speakers who are experienced with languages like French or English are likely to be aware that <CC> spellings represent underlying clusters. Such speakers are more likely to succeed in producing CC gestural sequence with a close transition, as in French or English consonant clusters. In this study, S3 and S6 had early exposure to languages with onset clusters, and insert acoustic vowels less frequently than the other speakers (S2, 4, 5, 7). These early bilinguals also have bimodal distributions of ICI durations. Diachronically, when these /CC/ loanwords originally entered Turkish from prestige languages like French, they were probably used by bilinguals who were highly conscious of their status as loanwords, and (variably) able to achieve a foreign-like gestural coordination. These bilinguals may even have been code-switching, and would have been aware that the borrowings began with /CC/, not /CVC/.

Second, hearing /CC/ loanwords in the speech of bilinguals, other speakers with less foreign language experience may recognize the underlying clusters, but fail to achieve a foreign-like gestural timing. This situation is comparable to English speakers producing

illegal onset clusters in Shaw & Davidson (2011). This would produce the gradient differentiation of non-lexical and lexical vowels found in the experiment as a whole, and exemplified in the data of S5 and S7, as well as S2.[5] These speakers could also be adapting an existing gestural coordination relation and its accompanying motor plan, perhaps one that governs the timing of onsets of adjacent syllables. Loanword phonology seeks to adapt a loanword to the existing phonological structures of the borrowing language, recycling native phonological processes to do so. We might expect a similar strategy of recycling native patterns at the level of gestural coordination as well. Crosslinguistically, this predicts that a language transitioning from a simpler syllable structure to a more complex syllable structure would exhibit vowel intrusion in the course of the transition, as speakers articulate complex new syllables by repurposing a limited set of existing gestural coordination plans that result in low overlap between consonant gestures.

Third, the presence of acoustic intrusive vowels in some tokens of complex onsets could result in some speakers reanalyzing the borrowed words as /CVC/. This occurred in the transcription task in Davidson (2007)—listeners sometimes transcribed [CᵊC] (containing a transitional schwa) as CVC. In Turkish, such reanalysis may be the source of orthographic alternations like *stil ~ sitil* 'style' and *klup ~ kulup* 'club.' Today, even Turkish monolinguals commonly use /CC/ loanwords, and might be expected to reanalyze intrusive vowels as underlying vowels. Anecdotally, Turkish children who are learning to write tend to write the intrusive vowel in onset clusters, and must be taught not to; this suggests that they are reanalyzing the words as starting with /CVC/ rather than /CC/. The prescriptive spelling of onset clusters in loanwords without an orthographic vowel probably works to maintain their representation as complex onsets—a hypothesis that should be investigated in future research.

Listeners who interpret the acoustic vowels they have heard as underlying vowels would not differentiate lexical and non-lexical vowels in their speech, either. This appears to describe the one monolingual speaker in this study, S4, and, to a lesser extent, her husband S2. Both S2 and S4 are from a smaller town in the province of Antalya, and exhibit a higher rate of acoustic insertion than speakers of the 'standard,' urban/Istanbul dialect (S3, 5, 6, 7) (**Figure 7**), as well as non-lexical vowels that are not significantly shorter than lexical vowels. While S2 exhibits F1 and F2 differences between lexical and non-lexical vowels (see Endnote 5), S4's non-lexical vowels do not differ significantly from lexical /ɯ/ in duration, F1, or F2.[6] This suggests that S4 may have reanalyzed the vowels in onset clusters as underlying /ɯ/. However, it cannot be that all S4's vowels in onset clusters result from reanalyzing those vowels as part of the underlying representations of familiar words, because S4 also produced acoustic vowels in novel nonce forms. That is, the insertion process generalizes beyond the conventionalized forms, even for speakers who do not acoustically differentiate lexical and non-lexical vowels. If onset cluster repair is not intrusion for S4, it must involve epenthesis, not only reanalysis.

To summarize, there is considerable interspeaker variation even in this small sample, which I suggest reflects variation among different speakers' grammars of gestural

---

[5] For S2, unlike S5 and S7, lexical status is not a significant predictor of vowel duration. For S2, lexical vowels had a mean duration of 39.55 ms, and non-lexical vowels had a mean duration of 37.11 ms. The difference was not significant ($t(462.03) = 1.34, p = 0.18$). But S2 does exhibit F1 and F2 differences between lexical and non-lexical vowels. Intrusive models outperform epenthetic models, whether harmony is assumed (F1: $\chi^2(3) = 40.62, p < 0.001$. F2: $\chi^2(3) = 42.04, p < 0.001$) or not (F1: $\chi^2(3) = 12.2, p < 0.01$. F2: $\chi^2(3) = 7.99, p < 0.05$).

[6] For S4, the mean duration of lexical vowels was 48.34 ms, and the mean duration of non-lexical vowels was 47.47 ms. This difference was not significant in a t-test ($t(472.82) = 0.34, p > 0.5$). Intrusive models of the duration and F1 of S4's vowels were also not significantly better than epenthetic models ($p > 0.05$), whether or not harmony was assumed. For F2, if harmony is assumed, lexical status is a significant predictor ($\chi^2(3) = 9.88, p = 0.01962$), but if no harmony is assumed, it is not ($\chi^2(3) = 2.84, p = 0.42$).

alignment. Some speakers apparently allow complex onsets and often achieve a close /CC/ coordination that does not produce an intrusive vowel. Other speakers also seem to allow complex onsets but employ a different gestural timing with less /CC/ overlap, resulting in a gradient distinction between the lexical and non-lexical vowels. Finally, some speakers do not differentiate lexical and non-lexical vowels; their grammars employ a /CC/ coordination with even less overlap, possibly because they still prohibit complex onsets and require epenthesis.

These three production strategies—categorical differentiation, gradient differentiation, and complete neutralization—correspond to the three strategies employed by different speakers in producing epenthetic vowels in Lebanese Arabic. Some speakers differentiate categorically from lexical vowels, some gradiently, and some not at all (Gouskova & Hall, 2009; Hall, 2013). However, the Lebanese speakers in Hall (2013) who differentiate epenthetic and lexical vowels did so consistently across items and repetitions, which is not the case for speakers in this study. Also, in Lebanese Arabic, the interspeaker variation is unlikely to be tied to language background since loanwords are not involved. One plausible reason for these differences between Lebanese Arabic and Turkish is that the insertion processes are phonologized to different degrees in the different languages. Vowel insertion in Lebanese is true phonological epenthesis, with interspeaker variation either in the degree of neutralization (Gouskova & Hall, 2009), or in the degree of cross-dialect influence (Hall, 2013). Each speaker of Lebanese Arabic realizes their epenthetic vowels in a predictable way.

In contrast, vowel insertion in Turkish is intrusion, produced by a gestural alignment that may be phonologized to different degrees for different speakers. Each speaker of Turkish does not realize consonant clusters in a predictable way. The acoustic variation within speakers could reflect gradience and ambiguity in speakers' mental representations, as in Gradient Symbolic Computation (Smolensky, Goldrick, & Mathis, 2014; Smolensky & Goldrick, 2016). Mental representations could be ambiguous between /CC/ and /CVC/, or could be solidly /CC/ but ambiguous as to the specific gestural coordination between the consonant gestures. Alternately, within-speaker variation could reflect failure to consistently achieve a targeted coordination, or other phonetic factors like speech rate. These conclusions are necessarily tentative, however, since the number of speakers here is so small. A future investigation of the factors that shape this intra- and inter-speaker variability could shed additional light on the mental representation of onset clusters for Turkish speakers, with possible implications for our understanding of loanword adaptation and its diachronic stability.

Furthermore, a perceptual study of Turkish onset cluster repair could clarify whether Turkish speakers, particularly monolinguals, are able to distinguish lexical and non-lexical vowels. If Turkish speakers use the acoustic differences to identify intrusive vowels in complex onsets, that would be a point in favor of an analysis where the gestural coordination that produces vowel intrusion is in fact grammaticized in Turkish, and maintained through perceptual cues. A perceptual study would also shed light on the ways in which factors like language-specific phonetic knowledge and the acoustic similarity of the stimuli, which have been shown to affect English speakers' perception of vowel intrusion (Davidson, 2007; Davidson & Shaw, 2012), also predict cross-linguistic perception of illegal consonant sequences.

Additional File

The additional file for this article can be found as follows:

- **Appendix.** Models. DOI: https://doi.org/10.5334/labphon.112.s1

Acknowledgements

Competing Interests

The author has no competing interests to declare.

References

Bellik, J. To appear. Turkish onset cluster repair: An ultrasound study. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 43*.

Boersma, P., & Weenink, D. 2015. Praat: Doing phonetics by computer [Computer program]. Version 5.4.08. Retrieved January 2015 from: http://www.praat.org/.

Bokhari, H., Durmaz, M., & Washington, J. 2016. An acoustic analysis of vowel insertion at syllable edges in Turkish. *Slides presented at the 2nd Conference on Central Asian Languages and Linguistics*. University of Indiana, Bloomington.

Bradley, T. 2004. Gestural timing and rhotic variation in Spanish codas. In: Face, T. L. (ed.), *Laboratory Approaches to Spanish Phonology*, 197–224. Berlin: Mouton de Gruyter.

Broselow, E. 2015. The typology of position-quality interactions in loanword vowel insertion. Retrieved from: https://linguistics.stonybrook.edu/faculty/ellen.broselow/selected.publications.

Browman, C., & Goldstein, L. 1993. Dynamics and Articulatory Phonology. *Haskins Laboratories Status Report on Speech Research,* SR-113, 51–62.

Cho, T., & Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics, 27*, 207–229. DOI: https://doi.org/10.1006/jpho.1999.0094

Clements, G., & Sezer, E. 1982. Vowel and Consonant Disharmony in Turkish. In: van der Hulst, H., & Smith, N. (eds.), *The Structure of Phonological Representations*. Dodrecht: Foris.

Davidson, L. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics, 34*(1), 104–137. DOI: https://doi.org/10.1016/j.wocn.2005.03.004

Davidson, L. 2007. The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology, 24*, 261–286. DOI: https://doi.org/10.1017/S0952675707001200

Davidson, L. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics, 38*(2), 272–288. DOI: https://doi.org/10.1016/j.wocn.2010.01.001

Davidson, L., & Shaw, J. 2012. Sources of illusion in consonant cluster perception. *Journal of Phonetics, 40*, 234–248. DOI: https://doi.org/10.1016/j.wocn.2011.11.005

Davidson, L., & Stone, M. 2003. Epenthesis Versus Gestural Mistiming in Consonant Cluster Production: An Ultrasound Study. *WCCFL 22 Proceedings*, Garding, G., & Tsujimura, M. (eds.), 165–178. Somerville, MA: Cascadilla Press.

Fleischhacker, H. 2005. *Similarity in phonology: Evidence from reduplication and loan adaptation*. (Doctoral dissertation), University of California, Los Angelos. (Order No. 3208408). Available from Dissertations & Theses @ University of California; ProQuest

Dissertations & Theses A&I. (305033127). Retrieved from: https://search.proquest. com/docview/305033127?accountid = 14523.

Gafos, A. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory, 20*(2), 269–337. DOI: https://doi.org/10.1023/A:1014942312445

Göksel, A., & Kerslake, C. 2005. *Turkish: A comprehensive grammar.* New York, NY: Routledge.

Gouskova, M., & Hall, N. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In: Parker, S. (ed.), *Phonological argumentation: Essays on evidence and motivation*, 203–225. London: Equinox Publishing.

Hall, N. 2003. *Gestures and segments: Vowel intrusion as overlap* (Doctoral dissertation). University of Massachusetts, Amherst. Available from Proquest. AAI3110499. https:// scholarworks.umass.edu/dissertations/AAI3110499.

Hall, N., 2006. Cross-linguistic patterns of vowel intrusion. *Phonology, 23*, 387–429. Cambridge University Press. DOI: https://doi.org/10.1017/S0952675706000996

Hall, N. 2011. Vowel Epenthesis. In: *The Blackwell Companion to Phonology*, Oostendorp, M., Ewen, C. J., Hume, E., & Rice, K. (eds.). DOI: https://doi.org/10.1002/9781444335262. wbctp0067

Hall, N. 2013. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics, 41*(2), 133–143. DOI: https://doi.org/10.1016/j. wocn.2012.12.001

Hall, N., & Sue, E. 2018. Hocank (Winnebago) vowel epenthesis: A phonological re-examination in light of phonetic data. *15th Old World Conference on Phonology*, London, United Kingdom. Abstract retrieved from: http://www.ucl.ac.uk/~ucjtcwh/ OCP15/abstracts/OCP_15_paper_96.pdf.

Inkelas, S., Küntay, A., Sprouse, R., & Orghun, O. 2000. Turkish Electronic Living Lexicon (TELL). *Turkic Languages, 4*, 253–275.

Kaun, A. 1995. *The typology of rounding harmony: An Optimality Theoretic approach* (Doctoral dissertation). University of California, Los Angelos. (Order No. 9530008). Available from ProQuest Dissertations & Theses A&I. (304173330). Retrieved from: https://search.proquest.com/docview/304173330?accountid = 14523.

Kaun, A. 1999. Epenthesis-Driven Harmony in Turkish. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society: Special Session on Caucasian, Dravidian, and Turkic Linguistics (2000)*, 95–106.

Kiliç, M. H., & Öğüt, F. 2004. A high unrounded vowel in Turkish: Is it a central or back vowel? *Speech Communication, 43*, 143–154. DOI: https://doi.org/10.1016/j. specom.2004.03.001

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. 2016. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. Retrieved from: https://CRAN.R-project.org/package = lmerTest.

Lees, R. B. 1966. On the interpretation of a Turkish vowel alternation. *Anthropological Linguistics, 8*(9), 32–39.

Lewis, G. 1967. *Turkish Grammar*. Oxford: Oxford University Press.

Murdoch, D., & Chow, E. D. 2018. ellipse: Functions for Drawing Ellipses and Ellipse-Like Confidence Regions. R package version 0.4.1. Retrieved from: https://CRAN.R-project. org/package = ellipse.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: https://www.R-project. org/.

Schmeiser, B. 2009. An acoustic analysis of intrusive vowels in Guatemalan Spanish /rC/ clusters. *University of Pennsylvania Working Papers in Linguistics, 15*(1), 193–202.

Shaw, J., & Davidson, L. 2011. Perceptual similarity in input-output mappings: A computational/experimental study of non-native speech production. *Lingua, 121*, 1344–1358. DOI: https://doi.org/10.1016/j.lingua.2011.03.003

Smolensky, P., & Goldrick, M. 2016. Gradient symbolic representations in grammar: The case of French liaison. Ms. http://roa.rutgers.edu/content/article/files/1552_smolensky_1.pdf.

Smolensky, P., Goldrick, M., & Mathis, D. 2014. Optimization and quantization in Gradient Symbol Systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science, 38*, 1102–1138. DOI: https://doi.org/10.1111/cogs.12047

Steriade, D. 1990. Gestures and autosegments: Comments on Browman's and Goldstein's paper. In: Kingston, J., & Beckman, M. (eds.), *Papers in Laboratory Phonology II: Between the Grammar and Physics of Speech*, 382–397. Cambridge: Cambridge University Press.

Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustic Society of America, 88*, 97–100. DOI: https://doi.org/10.1121/1.399849

Underhill, R. 1986. Turkish. *Studies in Turkish Linguistics*, Slobin, D., & Zimmer, K. (eds.). Philedelphia: John Benjamin Publishing Company.

Warner, N., Jongman, A., Cutler, A., & Mücke, D. 2002. The phonological status of Dutch epenthetic schwa. *Phonology, 18*, 387–420. DOI: https://doi.org/10.1017/S0952675701004213

Yavaş, M. 1980. Some pilot experiments on Turkish vowel harmony. *Linguistics, 13*(3), 543–562. DOI: https://doi.org/10.1080/08351818009370510

Yıldız, Y. 2010. *Age Effects in the Acquisition of English Onset Clusters by Turkish Learners: An Optimality-Theoretic Approach*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Yip, J. 2013. *Phonetic Effects on the Timing of Gestural Coordination in Modern Greek Consonant Clusters* (Doctoral dissertation). University of Michigan, Ann Arbor. http://hdl.handle.net/2027.42/102475.