Epenthesis and beyond

Recent approaches to insertion in phonology and its interfaces

Edited by

Ji Yea Kim

Veronica Miatto

Andrija Petrović

Lori Repetti

Topics in Phonological Diversity 4

language science press

Topics in Phonological Diversity

Editors: Natalia Kuznetsova (Università Cattolica del Sacro Cuore, Cormac Anderson (University of Surrey), Shelece Easterday (University of Hawai’i, Mānoa)

In this series:

1. Tallman, Adam J. R., Sandra Auderset & Hiroto Uchihara (eds.). Constituency and convergence in the Americas.
2. Bracks, Christoph. Compound Intonation Units in Totoli: Postlexical prosody and the prosody-syntax interface.
3. Nikulin, Andrey & Javier Carol. Historical phonology of Mataguayan.
4. Kim, Ji Yea, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.). Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces.

ISSN (print): 2943-6559
ISSN (electronic): 2943-6540

Epenthesis and beyond

Recent approaches to insertion in phonology and its interfaces

Edited by

Ji Yea Kim

Veronica Miatto

Andrija Petrović

Lori Repetti

language science press

Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.). 2024. *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces* (Topics in Phonological Diversity 4). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/469

ISBN: 978-3-96110-493-2 (Digital)
978-3-98554-122-5 (Hardcover)

ISSN (print): 2943-6559
ISSN (electronic): 2943-6540
DOI: 10.5281/zenodo.14193821
Source code available from www.github.com/langsci/469
Errata: paperhive.org/documents/remote?type=langsci&id=469

Cover and concept of design: Ulrike Harbort
Typesetting: Sebastian Nordhoff
Proofreading: Amir Ghorbanpour, Christian Döhler, Cong Luan Tran, David Carrasco Coquillat, Elliott Pearl, Frederic Blum, Jean Nitzke, Jeroen van de Weijer, Katja Politt, Lea Brown, Ludger Paschen, Mary Ann Walter, Mengyi Zhu, Rebecca Madlener, Sebastian Nordhoff, Silvie Strauß
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XƎLATEX

Language Science Press
Scharnweberstraße 10
10247 Berlin, Germany
http://langsci-press.org
support@langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Contents

Chapter 1

Epenthesis and beyond: An overview

Ji Yea Kim[a], Veronica Miatto[b], Andrija Petrović[b] & Lori Repetti[b]

[a]Gyeongsang National University [b]Stony Brook University

1 Introduction

Epenthesis, or the insertion of a non-etymological segment, has been an object of linguistic inquiry for centuries. The specific terms used to refer to the insertion of a segment at the beginning of a word (*pro(s)thesis*), within a word (*anaptyxis* or *svarabhakti*), or the end of a word (*paragoge* or *epithesis*), reflect the study of these processes within the Sanskrit, Greek, and Latin traditions (see Kiparsky 2022, Sen 2022, Oniga & Re 2021, etc.). Recently, the notion of insertion of non-etymological material has been expanded to include patterns that are not transparently phonologically motivated, but are conditioned phonetically, morphologically, morphosyntactically, and lexically. In the most familiar cases, the trigger for insertion is phonological in nature – for example, to reduce complexity in syllable structure – and the quality of the inserted segment is also determined phonologically – a featurally simple or predictable segment is used. However, an increasing body of research identifies cases where the motivation for insertion and the choice of the inserted segment lie beyond phonology. We outline some of these studies below: section 2 considers phonetic "intrusions", section 3 reviews canonical phonological epenthesis, and section 4 looks at morphological and morphosyntactic interactions. This is followed in section 5 by an overview of the articles included in this volume. These articles are a selection of the papers presented at the virtual workshop "Epenthesis and Beyond", held at Stony Brook University in 2021.

Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti. 2024. Epenthesis and beyond: An overview. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 1–20. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264528

2 Phonetic considerations: intrusion

Not all insertions are borne out of the same processes. While some segments are inserted phonologically, some are considered phonetic artifacts (see Ohala 1974, Ali et al. 1979, Hall 2006, to name a few). Systematic accounts of the distinction between phonologically and phonetically inserted segments is fairly recent (see Hall 2004) and traditionally binary. However, in the last few years, some authors (Grice et al. 2018, Karlin 2021, Hutin et al. 2021, Hall 2024 [this volume]) have challenged a binary distinction, stating that the dichotomy between epenthetic and intrusive segments is not always clear-cut.

Some inserted consonant-like sounds appear to be the result of a co-articulatory process. For example, Ohala (1974: 359) claims that [t]-insertion in a word like *false* /fɔls/, resulting in [fɔlts], is the result of articulatory transitions. In fact, the author states that the articulatory points of contact between /l/ and /s/ are somewhat complementary, so the transition between the two sounds might cause complete stoppage of the airflow, resulting in a t-like sound. Likewise, Ali et al. (1979) note that a stop might appear in nasal-fricative clusters, such as *warmth* /waɹmθ/, which can be pronounced as [waɹmpθ], and that the stop-insertion is due to delays in articulatory transitions.

In general, the consensus from these authors on English consonant intrusion seems to be that the "intrusive consonants" that appear in consonant clusters are optional (not always inserted), but there are some constraints on the process. Ohala (1974) reports that intrusive consonants are less likely to appear if the following consonant is voiced, while Ali et al. (1979) note that the inserted stop tends to be homorganic with the cluster's place of articulation. Fourakis & Port (1986) find that intrusive and "intended" stops are phonetically different, as the first is significantly shorter than the second. Regarding the perception of intrusive consonants, Warner & Weber (2001) find that these inserted stops are perceived by listeners only about 50% of the time in nonce words, and they are perceived differently, as shown by longer reaction times with intrusive stops. Finally, Leo (1985) claims that the addition of an intrusive consonant leaves the original syllable structure unaltered.

Vowel insertion, and its phonological status, is a much more hotly debated topic, as it has significant consequences for syllable structure. Some inserted vowels, called "excrescent vowels" or "intrusive vowels", have been argued to be a phonetic artifact, as they lack target gestures and do not interact with phonological processes. While these kinds of vowels have been identified in studies as early as Matteson & Pike (1958) and Levin (1987), it is not until Hall (2004, 2006, 2011) that we have a systematic account of the difference between intrusive and epenthetic vowels. According to Hall (2006), intrusive vowels are distinguished from epenthetic vowels because (i) they have a schwa-like quality or are a copy of a neighboring vowel, (ii) they occur in heterorganic clusters where at least one consonant is voiced, (iii) are usually optional, and (iv) speakers might not be aware of the presence of this vowel. Moreover, (v) their function is not to repair an illicit syllabic structure, but they might in fact serve a perceptual function, and they do not constitute the nucleus of a syllable.

Epenthetic vowels, on the other hand, (a) have a fixed or copied quality, (b) they occur in marked clusters that might or might not be voiced, (c) their presence is obligatory, and (d) speakers are usually aware of them. Moreover, (e) they constitute the nucleus of a syllable and are inserted to repair an illicit syllabic structure (Hall 2006, 2024 [this volume], for a more complete review of their diagnostics). According to the author, intrusive vowels and epenthetic vowels are, therefore, fundamentally different in their nature, the former being the result of articulatory retiming, while the latter being the result of a repair mechanism. Although intrusive vowels look like vowels phonetically, they do not participate in phonological processes, and they cannot be syllable nuclei. Since Hall (2006), there has been a plethora of studies looking more closely at inserted vowels and exploring this concept further, both in well-studied and understudied languages (see for example Bellik 2019b, 2024 [this volume] for work on Turkish, Burke et al. 2019 on Lamkang, Cavirani 2015 on two Lunigiana dialects, Grice et al. 2015 on Tashlhiyt Berber, Heselwood et al. 2015 on Libyan Arabic, Karlin 2021 on Finnish, Lancien & Côté 2019 on Quebec French, Nogita 2011 on Japanese, Pariente 2010 on Sephardic Hebrew).

Hall's specific research and the abovementioned studies address the dichotomy between epenthetic and intrusive vowels in consonant clusters. The nature of an intrusive vowel is in fact tied to the fact that it occurs in a consonant cluster, since it is due to articulatory retiming. However, there have been studies that claim that intrusive-like vowels can be found in word-final position as well. One example is Cavirani's (2015) work on the Lunigiana dialects of Italy, where Pontremolese and Carrarese are claimed to have schwa-like vocalic releases after consonants in word-final position. Another example is Italian where a vowel-like element is found after consonant-final words. Although the analysis is not undisputed (see Bafile 2002, 2005, Broniś 2016 and Passino 2008, who treat this vowel as the nucleus of a syllable, and therefore epenthetic), some accounts treat this vowel as non-syllabic (see Hamann & Miatto 2024 [this volume], Miatto et al. 2019, Miatto 2020, Repetti 2012), and therefore much closer to intrusive than epenthetic vowels. This is due to their similarities with intrusive vowels in quality (highly variable, usually transcribed as schwa, which is not a phoneme in the language), optionality (the vowel may or may not be inserted), and awareness of insertion (speakers are not aware that they are producing it and do not identify it as syllabic) (see in particular Miatto 2020).

Recently, many authors have acknowledged that the distinction between epenthetic and intrusive vowels is not clear cut. For example, while Grice et al. (2018) recognize that word-final inserted vowels in Italian resemble intrusive vowels in their appearance and seem not to be syllable nuclei, they claim that the inserted vowels are influenced by phonological and intonational pressures, and they state that the dichotomy between epenthetic and intrusive vowels is not completely satisfactory. In the same vein, Karlin (2021) reports that Finnish inserted vowels that have been previously called epenthetic, are actually intrusive vowels that are becoming phonologized. Likewise, Hutin et al. (2021) claim that French word-final inserted schwas, while sharing many properties with intrusive vowels, are ultimately epenthetic. These very recent studies, as well as Hall (2024 [this volume]), all agree that the distinction between intrusive and epenthetic is not always clear.

3 Phonology and epenthesis: canonical epenthesis

Epenthesis has usually been studied as a phonological process, and focus has been on the properties of the structure that is repaired and on the quality of the inserted segment(s). The generalization appears to be that unmarked segments are inserted to improve well-formedness. This type of epenthesis is called phonological or canonical epenthesis.[1] In this volume, we use the terms "marked", "unmarked", and "markedness" in the ways they are usually employed in the phonological literature, although we are fully aware of their complexity and the controversy that surrounds them. A question that can be raised in relation to epenthesis is what structures and segments are unmarked, under the assumption that unmarked vowels and consonants are inserted to repair marked structures. For example, /ə/ is universally an unmarked vowel and inserted to break illicit consonant clusters (Davidson & Stone 2003). Similarly, pharyngeals, the least marked consonants, such as glottal stop, are inserted to repair onsetless syllables (Lombardi 2002). Both cases espouse the Emergence of the Unmarked (TETU; McCarthy & Prince 1994) and are related to phonological naturalness.

On the other hand, Haspelmath (2006) provides an overview of markedness in different subfields of linguistics, such as phonetics, phonology, morphology, and

[1]Epenthesis has been described in sign languages as a means of ensuring syllable well-formedness (Brentari 1990).

semantics and gives alternative approaches to phonological (un)markedness. For instance, unmarked segments are inserted because they are more frequent and thus more predictable (Hume 2004). Another perspective highlights the role of language change in accounting for the appearance of synchronically unexpected segments (Blevins 2004), which is to be discussed in more detail later in this section.

Epenthetic processes usually result in phonological structures that are more acceptable in a given language, and have been discussed in the context of TETU, within the framework of Optimality Theory (OT, Prince & Smolensky 2004). While OT is the most common contemporary framework used when analyzing epenthesis, Government Phonology (Kaye 1990, Kaye et al. 1990) and the related CVCV Phonology (Scheer 2004) frame epenthesis, or the lack thereof, as the ability of a certain language to govern empty nuclei. Within this approach, final consonants are not thought of as codas, but as onsets of empty nuclei (Harris & Gussmann 1998). Therefore, if a language that does not allow for word-final consonants borrows loanwords with word-final consonants, its inability to govern empty nuclei will make it so that the nucleus will be phonetically realized, hence "epenthesized". See, for example, Bafile (2001, 2002) for an analysis of Emilian and Florentine dialects within a Government Phonology Framework, and Passino (2008) for an analysis of word-final gemination and schwa insertion in Italian within a CVCV framework.

No matter the framework, vowel epenthesis has been analyzed as improving syllable structure by resolving clusters or avoiding coda consonants, and the study of vowel epenthesis has played a key role in our understanding of syllable structure (Broselow 1982, Itô 1989, Piggott 1995, just to name a few). Consonant epenthesis also illustrates this point: it has been argued that epenthesis between vowels (1a) and before a word-initial vowel (1b) is a means of filling a missing onset resulting in an optimal CV(C) syllable.

(1)	a.	Persian	/sekei/	[sekeʔi]	'coin'-INDF	(Moradi 2017)
		Koryak	/alaal/	[alaʔal]	'summer'	(Kurebito 2004)
		English	*druid*	[dɹuwɪd]		
		English	*fire*	[fajəɹ]		
	b.	Persian	/abru/	[ʔabru]	'eyebrow'	(Dehghan & Kambuziya 2012)
		Koryak	/ajatək/	[ʔajatək]	'to fall'	(Kenstowicz 1976, Lombardi 2002)
		English	*apple*	[ʔæpəl]		

The non-etymological segment can be analyzed as inserted by a rule (Zwicky 1972, Dinnsen 1980) or resulting from the interaction and relative ranking of constraints. The latter is encoded in OT, which is the most frequently used framework in the phonology literature for epenthesis, with DEP and MAX constraints that militate against insertion and deletion, respectively, in output forms in comparison to input forms. McCarthy & Prince (1995) conceive of these constraints as applying to segments, and Lombardi (1998), Krämer (2003), Uffmann (2024 [this volume]), and others extend to include insertion/deletion of features.

The target of the insertion process is a marked structure, and the quality of the epenthetic segment is often claimed to be unmarked. The "unmarked" quality of the inserted segment was analyzed as determined by language-specific rules in early generative work (Zwicky 1972, Dinnsen 1980), and more recently by general principles such as the following universal hierarchy based on Place within the OT framework: *Dorsal, *Labial » *Coronal » *Pharyngeal (Lombardi 2002). Lombardi's hierarchy accounts for the insertion of a glottal stop, the least marked consonant, and one that is frequently observed cross-linguistically; if the least marked consonant is not available in a language, the next least marked consonant is inserted. Another type of hierarchy is Uffmann's (2007a, 2007b), in which syllabic positions and sonority are taken into consideration, and the optimal segment in a particular position is selected by the relative ranking of key constraints such as Dep(feature) and *Multiple.

A perception-based account is proposed by Jun (2015, 2021), following Steriade's (2001, 2009) P-map hypothesis: n-insertion in Korean compounds (for example, /som-ipul/ [somnipul] 'cotton sheet') is due to the fact that [n] is the perceptually least marked consonant before the high front vocoids /i/ and /j/ (Jun 2021: 34). Alternatively, the inserted segment can be a copy of a nearby segment (Kitto & Lacy 1999) or "split" from an adjacent input segment (Staroverov 2014).

There are other approaches in addition to markedness-based ones. Historical explanations deal with cases in which the quality of the inserted segment is clearly not "unmarked" from a synchronic perspective. Evolutionary Phonology (Blevins 2004, 2008) in particular accounts for phonologically opaque phenomena such as the emergence of [x] in Land Dayak (Blevins 2008). This is explained diachronically by a series of phonological processes that have taken place over time.

We have just discussed the insertion of a segment for phonetic or phonological reasons. The inserted segment can, over time, become part of the lexical item itself, as in the case of the /b/ in French *trembler* 'tremble' < Latin *tremulare*. Furthermore, it can even become a morphological marker, e.g., the active imperfective past morpheme in Modern Greek which originated as an epenthetic segment (Joseph & Ralli 2021). In the next section, we explore other ways in which epenthesis interacts with the grammar of the language beyond phonetics and phonology.

4 Morphological and morphosyntactic interactions

While there has been general agreement on the properties of canonical epenthesis, there are cases in which phonetics and phonology alone cannot account for the phenomena. Żygis (2010) provides an overview of epenthetic phenomena, focusing on consonant insertion, and notes that, in addition to canonical epenthesis, there exists a category of "grammatical insertions", consisting of morphologically, syntactically, and morphosyntactically conditioned insertion processes. Her paper provides a typology of consonantal insertions and reviews previous treatments, and she concludes that there exist very different analyses of consonant insertion because the processes that they model are fundamentally different.

Staroverov (2014) similarly makes a distinction between phonological epenthesis (which in his work is a result of Splitting, the operation that draws a correspondence between one input segment and multiple output segments), and morphologically restricted consonant-zero alternations. This is directly related to the difference between phonological epenthetic segments (which are predictable; in Staroverov's work, they share features with segments directly surrounding them) and morphologically restricted insertions (which are not predictable, so they permit a greater variety of segments).

Recent studies have focused on the factors influencing these other types of insertions. In some cases of conditioned epenthesis, the result is a phonologically improved structure, and one way to account for these insertions is to represent such segments as "ghosts" which are part of the underlying form and surface only when their presence is phonologically optimizing, as in French liaison: [le z̠ ami] 'the friends' vs. [le _ tami] 'the sieves'.[2] This is a purely phonological solution, circumventing the need for reference to morphosyntactic conditions, thus allowing for a strictly modular view of phonology (see, for instance, the analysis of Italian articles in Faust et al. 2018). Also, since these segments are an example of deletion but not insertion, especially when they disappear on the surface, the quality of the segments is therefore explicable in a diachronic aspect (i.e., not something that is inserted synchronically). Other work, on the other hand, defends the necessity of capturing the conditioning on the processes. In Zimmermann (2019), French liaison segments are examples of "appearing ghosts"; for the same

[2]"Liaison" segments are segments that were deleted historically, and their quality is therefore diachronically explicable.

phenomenon, Fukazawa (1999) and Pater (2010) employ lexically-indexed constraints, while Inkelas & Zoll (2007) argue for a co-phonology approach. Other analyses use mechanisms that more overtly refer to morphological structure in the input. For example, alignment constraints proposed by Jun (2015) and Blaylock (2017) penalize misalignment of morphological structure and phonological structure.

Cases of morphosyntactically conditioned epenthesis that result in phonologically more complex structures are surprisingly widespread. For example, Korean [n]-epenthesis results in the creation of a coda; crucially, [n]-epenthesis marks a morpheme boundary and is exclusive to compounds. Korean sometimes epenthesizes an [s] with the conjunctive suffix *-iraŋ* (for example, /pap-iraŋ/ [papsiraŋ] 'rice and'; Kim 2022b). Similarly, we find [s]-epenthesis in diminutives in Spanish (for example, /amoɾ-it-o/ [amoɾsito] 'love-DIM'; Kim 2022b) and English (for example, *Betsy* [bɛtsi] 'Elizabeth-DIM'; Kim 2022a,b). In each of these cases, a closed syllable is created, but syllable-morpheme alignment is optimized.

The quality of the non-canonical epenthetic segment can be determined by a variety of factors: [n] (used in Korean compounds) for perceptual markedness constraints (Jun 2015, 2021), [s] (in Korean *-iraŋ* suffixation, and Spanish and English diminutives) for frequency-based or analogical reasons (Kim 2022b), [o] (in final position in some Romance processes) since it is the morphologically neutral vowel (Aronoff & Repetti 2021), and other segments that are determined historically or lexically (Moradi et al. 2023). We also find cases of copy epenthesis at morpheme boundaries: the Korean innovative suffix *-lʌ* (/sʌul-ʌ/ [sʌullʌ] 'a person from Seoul') is formed by adding a copy of the final consonant of the stem *sʌul* 'Seoul' to the English agentive suffix *-er* (Kim 2022c), and examples of copy epenthesis at a word boundary are observed in many languages: Italian (/tram-elɛttriko/ [tram.me.lɛt.tri.ko] 'electric tram'), Jeju Korean (/kacuk-os/ [ka.cuk.kot] 'leather clothes') (Kim 2022b).

Finally, within the domain of syntax, the realization of a functional head with "default" phonological material, such as schwa as in the northern Italian dialect of Donceto [(ə) beːv] 'I drink', has been referred to as "syntactic epenthesis" (Cardinaletti & Repetti 2004).

Recent efforts have been made to provide a uniform account of the diverse epenthesis/insertion processes that have been observed. Moradi (2017) provides an overview of conditioned insertion processes in various languages, which have received different treatments, unifying them under the umbrella of non-canonical epenthesis and identifying the generalizations that characterize all of these phenomena. Aronoff & Repetti (2021) extend this survey to related processes in many Romance varieties. Petrović (2023) proposes a non-canonical epenthesis treatment for a pattern in Serbo-Croatian noun inflection, and provides a formalization couched in Boolean Monadic Recursive Schemes (BMRS; Chandlee & Jardine 2021). This work abstracts away from more common approaches partly due to the need to capture the fact that the epenthetic process under consideration is not necessarily phonologically optimizing. The formalism allows for direct reference to the input as well as the output forms, while at the same time the system does not surpass the computational complexity of phonological processes. (See also Moradi et al. 2023.)

5 Outline of the volume

The virtual workshop "Epenthesis and Beyond", held at Stony Brook University September 17–19, 2021, provided a forum for scholars who approach epenthesis and other types of insertion from new perspectives. The Workshop featured five invited speakers, fourteen 20-minute talks, and nine 5-minute blitz talks which substituted a poster session (a format more suitable for a virtual conference, held over Zoom).[3] The twelve articles included in this volume, which we summarize below, were selected from the many excellent papers that were presented at the Workshop, and they represent exciting new approaches to epenthesis.

5.1 Insertion or deletion?

Any analysis of a phenomenon as epenthesis must, of course, be superior to a competing deletion analysis of the same alternation. The differentiation of deletion and insertion processes is a topic that has sparked substantial discussion in the literature (e.g., Morley 2015), and even well-known and frequently referenced examples of epenthesis may merit a preferable deletion analysis (e.g., Staroverov 2015 on Ajyíninka Apurucayali, also pejoratively referred to as Axininka Campa). The chapters in Part I focus on such issues, and compare and evaluate competing hypotheses.

Words that are variably realized as CVCV or CCV can be analyzed as cases of optional epenthesis (CCV > CVCV) or optional deletion (CVCV > CCV). Hannah Sande investigates such alternations in her contribution "Insertion or deletion? CVCV/CCV alternations in Kru languages" (ch. 2). She offers diagnostics to determine if the alternation is best characterized as deletion or insertion, and concludes that in some Kru languages the alternation is the result of optional deletion (Guébie), while in others it is due to optional insertion (Dida) which

[3]Details of the Workshop can be found at https://www.stonybrook.edu/epenthesis/.

Sande characterizes as vowel intrusion rather than epenthesis. She concludes the article with proposals on the diachronic development of these CVCV/CCV alternations within a broader areal context.

The contribution by John Mansfield, Rosey Billington, and Hywel Stoakes, "Vowel predictability and omission in Anindilyakwa" (ch. 3), investigates vowels in the Australian language Anindilyakwa that are fully predictable (/a/ in word-final position) and highly predictable (non-low vowels in word-internal position), and can therefore be considered epenthetic. The authors find that the predictable vowels are frequently omitted in speech, while the unpredictable ones are never omitted; their investigation of a written wordlist is consistent with these findings. They adopt an information-theoretic approach to vowel predictability, which contributes to our understanding of segmental predictability and deletion.

5.2 Quality of epenthetic vowels

Another dimension to note when it comes to epenthesis is the quality of the inserted segments. In section 3, we discussed the cross-linguistically "unmarked" nature of some epenthetic segments (e.g., Zwicky 1972, Dinnsen 1980, Lombardi 2002). However, not all epenthetic segments can be explained by markedness, since there may be more than one epenthetic segment, and more importantly, not all epenthetic segments can be considered "unmarked". This leads our attention to other new approaches to epenthetic segment quality that are elaborated on in the following three articles.

Hassan Bokhari's paper, "The patterning of epenthesis in Urban Hijazi Arabic" (ch. 4), explores epenthesis in a variety of Arabic spoken in Saudi Arabia which has two different types of epenthesis using various epenthetic segments. "Syllable-structure-driven epenthesis" repairs a word-internal illegal string (such as *CVCCCV or *CV:CCV) with epenthetic [a], and "sonority-driven epenthesis" inserts [i], [u], or [a] to repair a word-final consonant cluster of rising sonority. In the latter case, the quality of the inserted vowel is determined by the stem vowel or the place feature of one of the consonants of the cluster.

Edward Rubin and Aaron Kaplan ("Segmental and prosodic influences on Bolognese epenthesis", ch. 5) report on epenthesis processes in a Romance variety spoken in northern Italy, Bolognese, that also involve more than one epenthetic vowel: [e] is the default segment, and [u] is used before [m] and [v]. Rubin and Kaplan propose that [v] is best treated as a sonorant since it patterns with [m] rather than the labial obstruents [p b f] in epenthesis processes, and since it alternates with [w]. In addition, they note that some obstruent-final clusters that are permitted word-internally, are instead repaired across word boundaries (such as verb-enclitic sequences). Their treatment of these latter cases as epenthesis rather than C/VC clitic allomorphy simplifies the analysis of clitics and unifies it with the analysis of epenthesis.

The article by Christian Uffmann, "Epenthesis as a matter of Faith" (ch. 6), provides an Optimality Theoretic account of the quality of default epenthetic segments. Cross-linguistically, certain epenthetic consonants (such as glottal stop) and vowels (such as schwa) are very common, and this fact has been accounted for by identifying these segments as "unmarked", violating few markedness constraints. Instead, Uffmann proposes an analysis based on Faithfulness. These particular segments are optimal because they violate few Faithfulness constraints, i.e., few violations of Dep(F): schwa is featureless, and glottal stop has laryngeal but not oral features.

5.3 Phonetics-phonology interface

Hall's (2004, 2006) seminal work on the distinction between epenthetic and intrusive vowels has sparked work that strives to determine the phonological status of inserted vowels in different languages, or that challenges the epenthetic-intrusive vowel dichotomy as too restrictive. The articles in this section address such questions with different approaches and frameworks, contributing to our understanding of insertion at the interface of phonology and phonetics.

Epenthesis strategies within Turkish (Turkic language family) onset and coda clusters differ. Vowel insertion in coda clusters is obligatory, reflected in the orthography, and invariant, and the inserted vowel harmonizes with the preceding lexical vowel. Alternatively, insertion in onset clusters is variable, not reflected in the orthography , and optional, and the inserted vowel is intermediate between a schwa-like vowel and a copy vowel. Using findings from an acoustic study (Bellik 2018, 2019a,b) and an ultrasound study inspired by Davidson & Stone (2003), Jennifer Bellik ("Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study", ch. 7) proposes that coda epenthesis is actual vowel insertion driven by syllable structure constraints, while insertion in onsets should be considered intrusion due to gestural alignment since the inserted vowel differs acoustically from lexical vowels and does not undergo harmony.

Nancy Hall ("Intrusive and epenthetic vowels revisited", ch. 8) reviews recent work on the distinction between epenthetic (phonologically active) vowels and intrusive (phonologically invisible) vowels, as well as subclasses of each. The main distinguishing feature between the two groups is the presence vs. absence of a new vowel gesture, respectively, but the typology of vowel insertion processes is actually much richer, and other factors need to be considered in order to characterize the full range of insertions.

Silke Hamann and Veronica Miatto expand the typology of intrusive vowels in their paper "Three language-specific phonological interpretations of release bursts and short vowel-like formants" (ch. 9). They show how the same phonetic material (word-final release burst) can have different phonological interpretations in different languages. In American English, a word-final release burst is interpreted as a plosive, and in Korean, as a vowel, despite the lack of vowel-like formants. In Italian, release bursts are followed by vowel-like formants which, however, are not perceived as vowels, making them more akin to intrusive vowels than epenthetic vowels. Hamann and Miatto adopt the Bidirectional Phonetics and Phonology model which distinguishes three levels of phonetic and phonological representations to account for these patterns.

Martin Krämer ("Prokaryotic syllables and excrescent vowels in two Yuman languages", ch. 10) also discusses the typology of intrusive segments in his investigation of two Yuman languages, Cocopa and Jamul Tiipay (Northwest of Mexico and Southwest of the United States). He analyzes certain structures in these languages as syllables consisting of a consonant or consonants (onset and optional coda) but without a nucleus or a mora, rendering them prosodically invisible. These types of structures have been referred to as degenerate, minor, or semisyllables, but he borrows the term "prokaryote" from biology to propose a new name for these nucleus-less syllables which provide evidence for another type of intrusive vowel.

5.4 Epenthesis and beyond

The ultimate goal of our workshop and of this collection of articles is to explore different kinds of insertion and different methods for analyzing epenthesis. The following articles are perhaps the embodiment of our overarching theme. They in fact complete our collection with analyses of epenthesis on understudied languages, using diverse methods and challenging the concept itself of insertion and epenthesis.

As the title of Michael Ramsammy's paper suggests, "On the diachrony of lateral epenthesis" (ch. 11) investigates the historical development of a particular type of consonant epenthesis in various languages including English, Motu (Oceanic language of Papua New Guinea), and Hindi. Ramsammy focuses on Hindi /l/-causatives, arguing against analyses involving allophony and analogy, and in support of an analysis whereby epenthetic /j/ changed to /l/ due to sonority optimization.

In their contribution "Textsetting the case for epenthesis in Armenian" (ch. 12), Luc Baronian and Nicolas Royer-Artuso show how textsetting, or the study of the way poets map their text onto a metrical grid, can be used as a tool to better understand the synchronic status of phonological processes such as schwa epenthesis. They use Armenian (Indo-European language family) as a test case, and build on the claim in Baronian (2017) that some cases of Western Armenian schwa are historically derived from epenthesis but are now part of the underlying form, and question whether schwa epenthesis is still productive. They map the lyrics of a song to the beats of the song, and conclude that schwa epenthesis is indeed used productively.

Brett Nelson's contribution, "Insertion of [spread glottis] at the right edge of words in Kaqchikel" (ch. 13), explores allophonic alternations in non-final and final position in Kaqchikel (a Mayan language spoken in Guatemala): plain stops are realized as aspirated, and non-nasal sonorants are realized as fricatives in final position; other consonants (glottalized stops, fricatives, nasals) do not alternate. Nelson proposes that the feature [+spread glottis] is inserted at the end of the word to mark the prosodic boundary. As a result, plain stops with [+sg] added become aspirated, and continuant (non-nasal) sonorants become obstruents (i.e., fricatives). The other consonants do not alternate because [+sg] would not change the segment (fricatives), or because [+sg] is blocked in certain contexts (glottalized stops, nasals). Nelson highlights the fact that even though [spread glottis] is not a contrastive feature in Kaqchikel, it plays an active role in its phonology.

6 Conclusion

The study of epenthesis is continuing to develop in significant ways; as we move forward, it remains a crucial area of research in linguistics. This introductory chapter has provided an overview of the past, present, and future of the study of epenthetic processes, highlighting the different types of epenthesis, including phonetic intrusion, canonical phonological epenthesis, morphological and morphosyntactic epenthesis, and the challenges and questions that arise from each.

The volume is organized into four parts, each of which explores different aspects of epenthesis. Part 1 focuses on the relationship between insertion and deletion processes, while Part 2 focuses on the quality of epenthetic segments. In Part 3, our attention turns to the phonetics-phonology interface and the interaction between the two domains. Finally, Part 4 expands the scope of epenthesis research by investigating understudied languages, employing diverse methodologies, and challenging the fundamental notions of epenthesis itself. These articles illustrate the complexity and richness of epenthesis, and its continued significance to phonological theory.

Overall, this volume seeks to contribute to our understanding of epenthesis and its role in language and linguistic theory, and we hope it will be a valuable resource for scholars interested in this area of research.

Acknowledgements

We would like to express our sincere thanks to participants of the Workshop "Epenthesis and Beyond", the reviewers of the talk abstracts and Proceedings papers, the Department of Linguistics at Stony Brook University, and in particular Alex Yeung for his technical support, and the Offices of the Provost, Vice President for Research, and the Dean of the College of Arts and Sciences at Stony Brook University for providing support for this Workshop.

References

Ali, Latif, Ray Daniloff & Robert Hammarberg. 1979. Intrusive stops in nasal-fricative clusters: An aerodynamic and acoustic investigation. *Phonetica* 36. 85–97.

Aronoff, Mark & Lori Repetti. 2021. Conditioned epenthesis in Romance. In Adam Ledgeway, John Charles Smith & Nigel Vincent (eds.), *Periphrasis and inflexion in diachrony*, 362–380. Oxford: Oxford University Press.

Bafile, Laura. 2001. La struttura vuota in fonologia: Il caso della sincope e dell'epentesi nei dialetti emiliani. *Annali dell'Università di Ferrara, Sezione di Lettere, Nuova Serie* 2. 147–167.

Bafile, Laura. 2002. Il trattamento delle consonanti finali nel fiorentino: Aspetti fonetici. In *Atti delle XXIII Giornate G. F. S. November 28-30*, 205–212. Pisa: ETS.

Bafile, Laura. 2005. Struttura sillabica e consonanti finali in varietà italiana. *Quaderni del Dipartimento di Linguistica - Università di Firenze* 15. 1–25.

Baronian, Luc. 2017. Two problems in Armenian phonology. *Language & Linguistics Compass* 11. e12247.

Bellik, Jennifer. 2018. An acoustic study of vowel intrusion in Turkish onset clusters. *Laboratory Phonology* 9(1). 1–16.

Bellik, Jennifer. 2019a. The effect of speech style and deaccentuation on vowel intrusion in Turkish complex onsets. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 3498–3502. Melbourne.

Bellik, Jennifer. 2019b. *Vowel intrusion in Turkish onset clusters*. University of California. (Doctoral dissertation).

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Blaylock, Reed. 2017. Driving derived environment effects with alignment constraints. *Proceedings of the Annual Meetings on Phonology* 4.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge & New York: Cambridge University Press.

Blevins, Juliette. 2008. Consonant epenthesis: Natural and unnatural histories. In Jeff Good (ed.), *Linguistic universals and language change*, 79–107. Oxford: Oxford Academic.

Brentari, Diane. 1990. Licensing in ASL handshape change. In Ceil Lucas (ed.), *Sign language research: Theoretical issues*, 57–68. Washington: Gallaudet University Press.

Broniś, Olga. 2016. Italian vowel paragoge in loanword adaptation. Phonological analysis of the Roman variety of Standard Italian. *Italian Journal of Linguistics* 28. 25–68.

Broselow, Ellen. 1982. On predicting the interaction of stress and epenthesis. *Glossa-an International Journal of Linguistics* 16. 115–131.

Burke, Mary, Shobhana Lakshmi Chelliah & Melissa Robinson. 2019. Excrescent vowels in Lamkang prefix sequences. *Journal of South Asian Languages and Linguistics* 6(2). 185–213.

Cardinaletti, Anna & Lori Repetti. 2004. Clitics in Northern Italian dialects: Phonology, syntax and microvariation. *University of Venice Working Papers in Linguistics* 14. 7–106.

Cavirani, Edoardo. 2015. *Modeling phonologization: Vowel reduction and epenthesis in Lunigiana dialects*. Utrecht: LOT.

Chandlee, Jane & Adam Jardine. 2021. Computational universals in linguistic theory: Using recursive programs for phonological analysis. *Language* 97. 485–519.

Davidson, Lisa & Maureen Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. *Proceedings of the West Coast Conference on Formal Linguistics* 22. 165–178.

Dehghan, Masoud & Aliyeh Kord-e Zafaranlu Kambuziya. 2012. A short analysis of insertion in Persian. *Theory and Practice in Language Studies* 2(1). 14–23. DOI: 10.4304/tpls.2.1.

Dinnsen, Daniel A. 1980. Phonological rules and phonetic explanation. *Journal of Linguistics* 16. 171–191.

Faust, Noam, Nicola Lampitelli & Shanti Ulfsbjorninn. 2018. Articles of Italian unite! Italian definite articles without allomorphy. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 63. 359–385.

Fourakis, Marios & Robert Port. 1986. *Stop epenthesis in English.* Journal of Phonetics 14. 197- 221.

Fukazawa, Haruka. 1999. *Theoretical implications of OCP effects on features in optimality theory*. PhD dissertation, University of Maryland.

Grice, Martine, Rachid Ridouane & Timo Roettger. 2015. Tonal association in Tashlhiyt Berber: Evidence from polar questions and contrastive statements. *Phonology* 32. 241–266. DOI: 10.1017/S0950147.

Grice, Martine, Michelina Savino & Timo B. Roettger. 2018. Word final schwa is driven by intonation: The case of Bari Italian. *The Journal of the Acoustical Society of America* 143(4). 2474–2486. DOI: 10.1121/1.5030923.

Hall, Nancy. 2004. *Vowel intrusion as overlap*. University of Massachusetts. (Doctoral dissertation).

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Hamann, Silke & Veronica Miatto. 2024. Three language-specific phonological interpretations of release bursts and short vowel-like formants. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 199–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264544.

Harris, John & Edmund Gussmann. 1998. Final codas: Why the west was wrong. In Eugeniusz Cyran (ed.), *Structure and interpretation: Studies in phonology*, 139–162. Lublin: Folium.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42. 25–70.

Heselwood, Barry, Abdurraouf Shitaw, Aimen Ghummed & Leendert Plug. 2015. Epenthetic and excrescent vowels in stop sequences in Tripolitanian Libyan

Arabic. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Glasgow.

Hume, Elizabeth. 2004. Deconstructing markedness. *Berkeley Linguistics Society* 30.

Hutin, Mathilde, Adèle Jatteau, Yaru Wu, Ioana Vasilescu, Lori Lamel & Marine Adda-Decker. 2021. A corpus-based study of the distribution of word-final schwa in Standard French and what it teaches us about its phonological status. *Isogloss. Open Journal of Romance Linguistics* 7. 1–27.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133–171.

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language & Linguistic Theory* 7. 217–259.

Joseph, Brian & Angela Ralli. 2021. *The aftermath of epenthesis: From hiatus to morpheme in Greek*. Paper presented at the Epenthesis Workshop, Stony Brook University, September 17-19, 2021.

Jun, Jongho. 2015. Korean n-insertion: A mismatch between data and learning. *Phonology* 32(3). 417–458. DOI: 10.1017/S0950275.

Jun, Jongho. 2021. Morphophonological gradience in Korean n-insertion. *Glossa* 6. 1–40. DOI: 10.5334/gjgl.1401.

Karlin, Robin. 2021. Finnish inserted vowels: A case of phonologized excrescence. *Nordic Journal of Linguistics* 45(1). 49–79.

Kaye, Jonathan. 1990. 'Coda' licensing. *Phonology* 7(2). 301–330.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7(2). 193–231.

Kenstowicz, Michael. 1976. Some rules of Koryak phonology. *Studies in the Linguistic Sciences* 6. 22–37.

Kim, Ji Yea. 2022a. Consonant epenthesis in English diminutives. *Korean Journal of Linguistics* 47(3). 385–406. DOI: 10.18855/lisoko.2022.47.3.001.

Kim, Ji Yea. 2022b. *Morphophonological consonant epenthesis in suffix realizations*. PhD dissertation. Stony Brook University.

Kim, Ji Yea. 2022c. The Korean innovative suffix -lʌ: Indirect affix borrowing and morphological copy epenthesis. *Studies in Phonetics, Phonology and Morphology* 28(1). 57–82. DOI: 10.17959/sppm.2022.28.1.57.

Kiparsky, Paul. 2022. Pāṇini. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*, 1–24. Oxford: Oxford University Press.

Kitto, Catherine & Paul V. De Lacy. 1999. Correspondence and epenthetic quality. In Catherine Kitto & Carolyn Smallwood (eds.), *Proceedings of AFLA VI:*

The Sixth Meeting of the Austronesian Formal Linguistics Association, 181–200. Toronto: Department of Linguistics, University of Toronto.

Krämer, Martin. 2003. *Vowel harmony and correspondence theory*. Berlin: Mouton de Gruyter.

Kurebito, Megumi. 2004. A report on Koryak phonology. *Languages of the North Pacific Rim* 9. 117–144.

Lancien, Mélanie & Marie-Hélène Côté. 2019. Intrusive vowels preceding /r/ in Quebec French. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2015)*. Melbourne.

Leo, Wetzels. 1985. The historical phonology of intrusive stops: A non-linear description. *The Canadian Journal of Linguistics/Revue Canadienne De Linguistique* 30. 285–333.

Levin, Juliette. 1987. Between epenthetic and excrescent vowels. *Proceedings of the West Coast Conference on Formal Linguistics* 6. 187–201.

Lombardi, Linda. 1998. *Evidence for MaxFeature constraints from Japanese*. Unpublished manuscript, University of Maryland [Rutgers Optimality Archive, ROA-247].

Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19. 219–251.

Matteson, Esther & Kenneth L. Pike. 1958. Non-phonemic transition vocoids in Piro (Arawak). *Miscellanea Phonetica* 3. 22–30.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *University of massachusetts occasional papers in linguistics 18*, 249–384. Amherst: GLSA Publications [Rutgers Optimality Archive, ROA-103].

McCarthy, John J. & Alan S. Prince. 1994. *The emergence of the unmarked*. Ms. University of Massachusetts, Amherst and Rutgers University.

Miatto, Veronica. 2020. Perception of word-final inserted vowels and syllabicity in Italian. In *Proceedings of the 11th International Conference of Experimental Linguistics (ExLing 2020)*, 137–140.

Miatto, Veronica, Silke Hamann & Paul Boersma. 2019. Self-reported L2 input predicts phonetic variation in the adaptation of English final consonants into Italian. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*, 949–953. Melbourne: International Phonetic Association.

Moradi, Sedigheh. 2017. *Non-canonical Epenthesis: Epenthetic quality and the role of morphonology*. Ms, Stony Brook University. https://www.academia.edu/100511567/Non_canonical_Epenthesis_Epenthetic_quality_and_the_role_of_morphonology.

Moradi, Sedigheh, Andrija Petrović, Lori Repetti & Mark Aronoff. 2023. *Epenthesis goes beyond Phonology: A unified formal account*. Ms, Stony Brook University. https://github.com/andrija-petrovic/andrija-petrovic.github.io/blob/master/pdfs/Epenthesis_Goes_Beyond_Phonology.pdf.

Morley, Rebecca L. 2015. Deletion or epenthesis? On the falsifiability of phonological universals. *Lingua* 154. 1–26.

Nogita, Akitsugu. 2011. Epenthesis, intrusion, or deletion? Vowel alternation in consonant clusters by Japanese ESL learners. *Proceedings of the 27th Northwest Linguistics Conference* 21. 43–51.

Ohala, John. 1974. Experimental historical phonology. In John Mathieson Anderson & Charles Jones (eds.), *Historical linguistics II. Theory and description in phonology*, 353–389. Amsterdam: North-Holland Publishing Co.

Oniga, Renato & Alessandro Re. 2021. *Epenthesis in Latin*. Paper presented at the Epenthesis Workshop, Stony Brook University, September 17-19, 2021.

Pariente, Itsik. 2010. Pharyngeal related non-lexical vowels in Sephardic Modern Hebrew. *Linguistics in Amsterdam* 3. 1–19.

Passino, Diana. 2008. *Aspects of consonantal lengthening in Italian*. Padua: Uni-Press.

Pater, Joe. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.), *Phonological argumentation*, 123–154. London: Equinox.

Petrović, Andrija. 2023. A formal account of morphological epenthesis in Serbo-Croatian. *Morphology* 33(3). 335–359. DOI: 10.1007/s11525-023-09412-9.

Piggott, Glyne L. 1995. Epenthesis and syllable weight. *Natural Language & Linguistic Theory* 13. 283–326.

Prince, Alan & Paul Smolensky. 2004. *Optimality theory: Constraint interaction in Generative Grammar*. Cambridge: Blackwell.

Repetti, Lori. 2012. Consonant-final loanwords and epenthetic vowels in Italian. *Catalan Journal of Linguistics* 11. 167–188.

Scheer, Tobias. 2004. *What is CVCV and why should it be?* Berlin: De Gruyter Mouton. DOI: 10.1515/08336.

Sen, Ranjan. 2022. The Greco-Roman tradition. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*, 109–133. Oxford: Oxford University Press.

Staroverov, Peter. 2015. *[t]-epenthesis revisited: The case of Apurucayali*. Ms, Universität Leipzig. https://ling.auf.net/lingbuzz/002379/.

Staroverov, Petr. 2014. *Splitting theory and consonant epenthesis*. Rutgers University. (Doctoral dissertation). https://rucore.libraries.rutgers.edu/rutgers-lib/45498/.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 219–250. San Diego: Academic Press.

Steriade, Donca. 2009. The phonology of perceptibility effects: The p-map and its consequences for constraint organization. In Kristin Hanson & Sharon Inkelas (eds.), *On the nature of the word: Studies in honor of Paul Kiparsky*, 151–179. Cambridge: MIT Press.

Uffman, Christian. 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29. 451–476.

Uffmann, Christian. 2007. *Vowel epenthesis in loanword adaptation* (Linguistische Arbeiten 510). Tübingen: Niemeyer.

Uffmann, Christian. 2024. Epenthesis as a matter of FAITH. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 123–141. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264538.

Warner, Natasha & Andrea Weber. 2001. Perception of epenthetic stops. *Journal of Phonetics* 29. 53–87.

Zimmermann, Eva. 2019. Gradient symbolic representations and the typology of ghost segments. In *Proceedings of AMP 2018*.

Zwicky, Arnold. 1972. Note on a phonological hierarchy in English. In Robert Stockwell & Ronald Macauley (eds.), *Linguistic change and generative theory*, 275–301. Bloomington: Indiana University Press.

Żygis, Marzena. 2010. Typology of consonantal insertions. *ZAS Papers in Linguistics* 52. 111–140.

Chapter 2

Insertion or deletion? CVCV/CCV alternations in Kru languages

Hannah Sande
UC Berkeley

Kru languages, spoken in Liberia and Côte d'Ivoire, show an alternation between CVCV and CCV in a subset of lexical items. The alternation is variable, where the same word can appear as either CVCV or CCV in a single environment, and lexically specific, where only a subset of morphemes alternate in this way. This paper describes the CVCV/CCV alternation in Kru, showing that whether this pattern is best categorized as an instance of vowel insertion or deletion differs by language. Two representative Kru case studies are presented in detail: The CVCV/CCV alternation in Dida is shown to be best analyzed as underlying /CCV/ with epenthesis or vowel intrusion, whereas the CVCV/CCV alternation in Guébie is shown to involve underlying /CVCV/ sequences with optional deletion of the initial vowel. Diagnostics are proposed for determining when a vowel/∅ alternation is synchronically best analyzed as insertion or deletion. CVCV/CCV alternations found in nearby Kru, Mande, and Kwa languages are examined, leading to conclusions about the diachronic development of the CVCV/CCV alternation in West Africa.

1 Introduction

Kru languages, spoken in Liberia and Côte d'Ivoire, show a CVCV/CCV alternation, as in the Guébie example in (1).

(1) bala$^{3.3}$/bla^{3}
hit
'hit' (Guébie, Eastern Kru)

Hannah Sande. 2024. Insertion or deletion? CVCV/CCV alternations in Kru languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 21–55. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264530

In some Kru languages this alternation has been analyzed as vowel insertion, while in others it has been called vowel deletion. This paper aims to describe the CVCV/CCV alternation in multiple Kru languages, and to diagnose for each language whether the alternation is best analyzed as synchronic deletion or insertion. As will be shown in Sections 2 and 3, the alternation seems to be best analyzed as V-deletion in some Kru languages but V-insertion in others. The two case studies examined in detail come from Dida, presented in §2, and Guébie, in §3, both part of the Neyo-Dida group of Eastern Kru languages. §4 examines the wider picture of the CVCV/CCV alternation in Kru, Mande, and Kwa languages, drawing historical and areal conclusions about the diachronic path of this alternation. §5 summarizes the diagnostics proposed throughout the paper for determining when a vowel/∅ alternation is best analyzed as synchronic insertion versus deletion.

1.1 Background on CVCV/CCV alternations

1.1.1 V/∅ as insertion

/CCV/ → [CVCV] alternations are often analyzed as *copy epenthesis*, also called Dorsey's Law: A vowel is inserted in a CCV word to break up the CC cluster (Miner 1979, 1989, Hale & White Eagle 1980, Hayes 1995, Clements 1986, 1991, Halle 2000, Kawahara 2007, Stanton & Zukoff 2018). For a specific implementation of Dorsey's law involving feature spreading, see Kawahara (2007), and for an analysis involving correspondence and identity, see Stanton & Zukoff (2018). In such cases, the quality of the inserted vowel tends to match the quality of the underlying vowel, or surface as the predictable epenthetic vowel in the language. The inserted vowel is often shorter than an underlying vowel, and lacks its own prosody (e.g. no independent stress or tone).

(2) Winnebago /prás/ → [parás] (Miner 1979: 27)

CCV/CVCV alternations have also been analyzed as due to retiming of articulatory gestures, where an underlying CCV sequence is pronounced as CVCV due to gestural overlap (Hall 2003, 2006, 2011, Hall 2024 [this volume], Bellik 2024 [this volume]). For Hall, there is a distinction between phonologically epenthesized vowels repairing a marked structure (i.e., vowel epenthesis to break up a cluster as per Dorsey's Law), and intrusive vowels which are phonologically invisible and involve a rearrangement of articulatory targets. Hall (2006: 391) provides diagnostics for distinguishing between phonologically visible epenthetic vowels and phonologically invisible intrusive vowels (see also Hall 2024 [this volume]), repeated in (3) and (4), below.

(3) Properties of phonologically invisible inserted vowels (intrusive vowels) (Hall 2006: 391)
   a. The vowel's quality is either schwa, a copy of a nearby vowel, or influenced by the place of the surrounding consonants.
   b. If the vowel copies the quality of another vowel over an intervening consonant, that consonant is a sonorant or guttural.
   c. The vowel generally occurs in heterorganic clusters.
   d. The vowel is likely to be optional, have a highly variable duration or disappear at fast speech rates.
   e. The vowel does not seem to have the function of repairing illicit structures. The consonant clusters in which the vowel occurs may be less marked, in terms of sonority sequencing, than clusters which surface without vowel insertion in the same language.

(4) Properties of phonologically visible inserted vowels (epenthetic vowels) (Hall 2006: 391)
   a. The vowel's quality may be fixed or copied from a neighboring vowel. A fixed-quality epenthetic vowel does not have to be schwa.
   b. If the vowel's quality is copied, there are no restrictions as to which consonants may be copied over.
   c. The vowel's presence is not dependent on speech rate.
   d. The vowel repairs a structure that is marked, in the sense of being cross-linguistically rare. The same structure is also likely to be avoided by means of other processes within the same language.

Hall's diagnostics are used throughout this paper to distinguish between types of inserted vowels. Note that these diagnostics do not necessarily distinguish between vowel insertion versus deletion, but rather between multiple types of vowel insertion. A separate set of diagnostics for determining whether a given vowel is best analyzed as inserted, or underlying and subject to deletion, is developed throughout the paper and presented in §5.

1.1.2 V/∅ as deletion

Vowel/∅ alternations are also sometimes analyzed as deletion, primarily when the quality of the vowel that participates in the alternation is not predictable given the phonological context. Instances of vowel deletion tend to, but do not necessarily, interact with stress and syllable weight (Hawkins 1950, Anderson

1965, Al-Mozainy 1981, Willett 1982, Willett 1991, Fitzgerald 1997, 2002, Kager 1997, Alan 2000, Jacobs 2004, Riggle 2006, McCarthy 2008).

One example of a language with a Vowel/∅ alternation that has been analyzed as deletion comes from Southeastern Tepehuan (ST), (Willett 1982, Willett 1991, Kager 1997). In ST, all words have stress on the first or second syllable, whichever is heaviest (CVV ≫ CVC ≫ CV). If they are both equally heavy, the first syllable is stressed (Willett 1982: 176). Long vowels shorten and short vowels delete in unstressed syllables, which is particularly visible in reduplication contexts where the vowel shortens or deletes in one copy (the base) but not the other (the reduplicant). In reduplication the initial syllable is copied and prefixed to the base. If the initial syllable is underlyingly heavy as in (5a), the first syllable of the base is shortened in reduplication, resulting in a light unstressed syllable. If the initial syllable is light as in (5b), then the first vowel of the base, which is a short unstressed vowel, is deleted. We know that the vowel is underlyingly present because it is present in non-reduplicated forms, as well as in the reduplicant itself, but not in the base.

(5) Southeastern Tepehuan syncope

		Underlying	Surface	
a.	Singular	/kooʔ/	(kóoʔ)	‘snake’
	Plural	/koo-kooʔ/	(kóo).koʔ	‘snakes’
b.	Singular	/topaa/	(to.páa)	‘pestle’
	Plural	/to-topaa/	(tót.pa, *tó.to.pa)	‘pestles’

The quality of the vowel in the V/∅ alternation in ST is not predictable from the phonological context. Additionally, the result of deletion is phonologically optimizing: stressed syllables are heavy but unstressed syllables are light. These two facts lead to the conclusion that this alternation involves deletion, and that it is phonologically motivated and derived.

1.1.3 Diagnosing deletion versus insertion

Based on previous work on CVCV/CCV alternations, the following two lines of investigation can serve as a starting point in determining whether a CVCV/CCV alternation involves deletion or insertion. First, if the first vowel in CVCV forms that alternate with CCV is unpredictable given the phonological context, it must be analyzed as underlyingly present and deleted: CVCV→CCV. On the other hand, a purely phonologically predictable initial vowel in alternating CVCV/CCV forms points to underlying /CCV/ plus vowel insertion. Second, we can determine whether the first vowel in an alternating CVCV form participates independently in the phonology of the language from the second vowel: Can the first vowel host its own tone or stress? Can it participate in phonological processes independently from the second vowel? If the first vowel is prosodically and phonologically independent of the second vowel, it is likely not intrusive, and may involve deletion or, if its quality is phonologically predictable, it may be analyzable a phonologically visible epenthetic vowel (see §1.1.1 for Hall's diagnostics for intrusive versus epenthetic vowels).

1.2 Background on Kru languages

Kru languages are spoken in Liberia and Côte d'Ivoire (Figure 1).

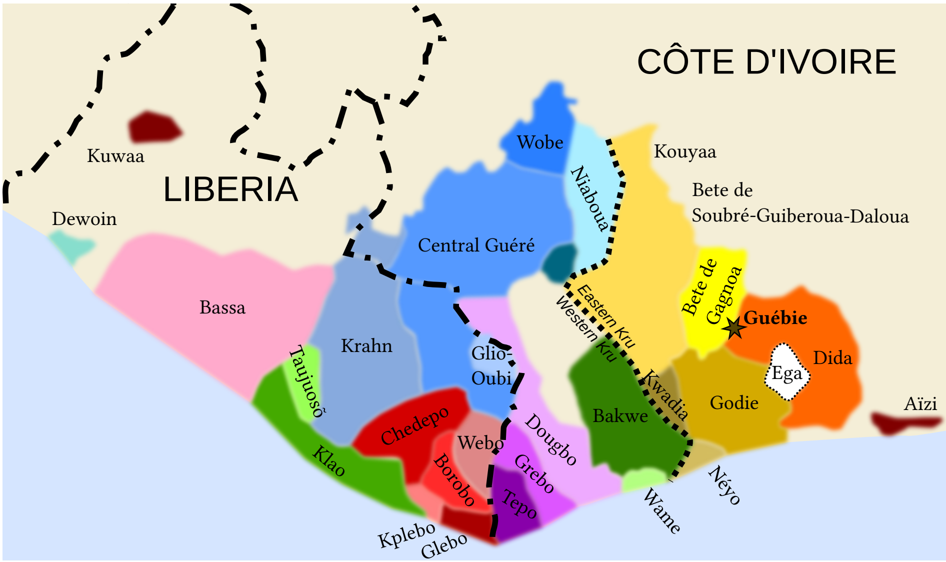


Figure 1: Kru languages (Marchese 1983)

The Kru language family tree in Figure 2 shows a split between Eastern and Western Kru, for which there is very clear historical evidence (Marchese 1986, Zogbo 2012, 2019), as well as widely agreed-upon groups within Eastern and Western Kru. Not included are Aizi, Kuwaa, and Seme, which are sometimes called Kru isolates, but whose classification as Kru at all is debated (Marchese 1983, Hérault 1971, Vogler 2015). The languages examined in depth in this paper come from the Neyo-Dida group, in bold.

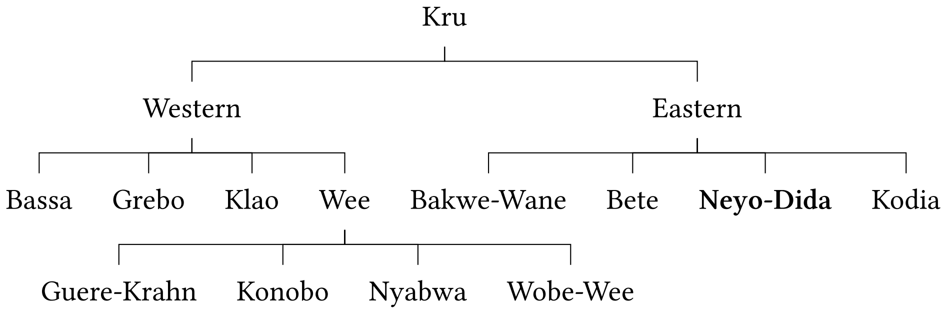


Figure 2: Kru language family tree

There are several canonical phonological properties of Kru languages, including contrastive labiovelar and labialized velar stops in addition to voiced and voiceless stops at other places of articulation. There is a bilabial implosive in all Eastern Kru languages and many Western Kru languages (not so in so-called "Kru isolates", Marchese 1983: 41), as well historical evidence for a Proto-Kru alveolar implosive (Zogbo 2012). In many Western Kru languages with contrastive nasal vowels, nasal consonants are analyzed as non-contrastive. Rather, they are allophones of oral consonants that surface adjacent to a nasal vowel (e.g. /l/ → [n] / _[+nasal]), (Zogbo 2019). Many Eastern Kru languages show synchronic alternations between sonorant consonants and nasal consonants in the environment of a nasal consonant, though there are also demonstrably contrastive nasal consonants and a lack of nasal vowels in Eastern Kru. All Kru languages show vowel harmony, often ATR harmony, and sometimes also height harmony (Marchese 1983).

Kru languages have multiple contrastive tone heights, tones that differentiate lexical items, and grammatical tone patterns. No Kru language has fewer than three contrastive tone heights, and many have four. There are also contour tones on single tone-bearing units that can be analyzed as made up of sequences of level tones. Grammatical tone marks aspect distinctions, negation, and nominative versus accusative versus genitive case. Throughout, tone is marked as in the original data source, sometimes with numeral superscripts where higher numbers represent higher tones, and sometimes with diacritics on vowels.

Syllables in Kru languages are almost exclusively CV, though V-initial syllables are possible in pronouns and loan words. There are complex segments such as [kp, gb, k^w, g^w], though there are no consonant clusters other than the alternating CVCV/CCV sequences that are the topic of this paper.

2 Insertion in Dida

2.1 Language background

Dida is a dialect cluster spoken in south-central Côte d'Ivoire. This section focuses on the Lakota cluster of Dida languages, which itself is made up of three varieties: Lakota, Abou, and Vata. Dida Lakota is spoken in the town of Lakota and surrounding village communities by about 93,800 people as of 1993, according to Ethnologue (Eberhard et al. 2020).

The data presented here comes from existing descriptions of Dida in the literature: Guéhoun (1993) on Dida Lakota, Vogler (1976), Kaye (1981), Kaye & Charette (1981), Kaye (1982) on Dida Lakota and Vata, and Masson (1992) on Dida Yocoboué.

The vowel inventory of Dida is provided in Figure 3.

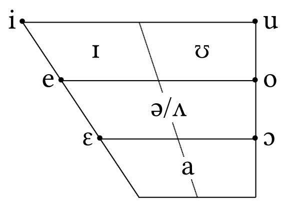


Figure 3: Dida Vowel inventory

The consonant inventory is provided in Table 1.

Table 1: Dida consonant inventory (Guéhoun 1993: 38)

	Bilabial		Lab. dent.		Alveo-pal.		Palatal		Velar		Lab. vel.	
Plosive	p	b			t	d	c	ɟ	k	g	kp	gb
Nasal		m				n		ɲ		ŋ		
Fricative			f	v	s	z						
Approx		ɓ				l		j				w

The Vata variety is analyzed as having four distinct tone heights (Kaye & Charette 1981), while other Dida varieties have three contrastive tone heights (Guéhoun 1993: 68), written with diacritics on vowels (the diacritic on mid-toned vowels is often left off). Grammatical tone differentiates aspect, negation, and case, as in other Kru languages.

Guéhoun (1993) analyzes /g^w, k^w, $ŋ^w$/ as sequences of two consonants, but they distributionally function as singletons, and are analyzed as single segments in all other sources on Dida and other Kru languages (cf. Marchese 1983: 39). They are treated as singletons here. Otherwise, the only surface clusters in the language are found in alternating CVCV/CCV forms.

2.2 Insertion or deletion?

In Dida, all alternating CVCV/CCV forms share the property of having /l/ as their second consonant, CVLV/CLV (though note that /l/ surfaces as [r] after certain consonants in Dida, as in Table 2c). Additionally, the first vowel in alternating CVCV forms always matches the vowel quality of the following vowel, and the tone on the first vowel always exactly matches that of the second vowel.

In Table 2 the alternating sequences are underlined. Note that the alternating CVCV sequence can be the first or last CVCV sequence of a surface CVCVCV form (Table 2a,d,e), or it can be the entirety of a surface CVCV form (Table 2b,c).

Table 2: Dida Lakota CVCV ~ CCV alternations (Guéhoun 1993: 56)

	CVCV	CCV	
a.	wʊ̀lʊ̀lɪ	wlʊ̀lɪ	‘to leave’
b.	ŋɛlɛ	ŋlɛ	‘smell’
c.	ɟulu	ɟru	‘salt’
d.	kpokele	kpokle	‘stool/chair’
e.	dugbulu	dugblu	‘village center’

The synchronic facts in Dida point to the first vowel (V1) in alternating CVCV sequences as being inserted, rather than deleted in CCV sequences. All cases of alternating CVCV ~ CCV sequences in Dida exhibit a V1 identical in features to V2, which means the initial vowel is entirely phonologically predictable. Additionally, all cases exhibit identical tone on V1 and V2, with no evidence that the first vowel displays independent prosodic or phonological behavior from the second vowel. Outside of these alternating forms, sequences of distinct tones on non-alternating CVCV sequences are quite common (e.g. palɛ̀li ‘enter’ with MLM), as are sequences of CVCV with non-identical vowels where the second consonant is /l/ (e.g. wɔlʊ ‘granary’).

Additional evidence for a vowel-insertion analysis of Dida CVCV/CCV sequences comes from the surface tone patterns in words longer than two syllables, where the alternating CVCV/CCV is initial. Specifically, in Dida, contour tones and sequences of identical level tones only surface at the right edge of a word. Tones in Dida can be analyzed as associating to vowels (or tone-bearing units) one-to-one from left-to-right, such that any one-to-many (adjacent identical tone sequences) or many-to-one (contour tones) tone-to-vowel mappings occur at the right edge of a word, and not the left (Guéhoun 1993). The L.L.M (low, low, mid) tonal pattern in Table 2a is not a possible lexical tone melody in Dida; it is unattested outside of alternating forms with an initial CVCV/CCV sequence. This initial one-to-many mapping, where two adjacent syllables surface with the same level tone, is unexpected given the regular tone association process in Dida; typically, on a three-syllable word with a L.M melody the Dida, the tone-to-vowel mapping results in a surface L.M.M pattern. The tonal pattern in Table 2a would thus need to be analyzed as an exception to the otherwise robust left-to-right tone-to-vowel association patterns in Dida, unless the initial vowel in [wʊ̀lʊ̀lɪ] is inserted after tonal association takes place. The tonal patterns provide an additional piece of evidence that alternating CVCV/CCV sequences in Dida are best analyzed as /CCV/ with vowel insertion.

We can now consider whether the inserted vowels in Dida are phonologically visible epenthetic vowels, or phonologically invisible intrusive vowels, based on the diagnostics from Hall (2006) given in (3) and (4). The initial vowel in alternating CVCV sequences in Dida has the same features as the following vowel. This type of vowel feature copying does not, on its own, distinguish between being epenthetic (4a) or intrusive (3a). However, the intervening consonant is always a sonorant (/l/), as predicted for cases of vowel intrusion (3b), and the vowel is optional and dependent on speech rate, as predicted for cases of vowel intrusion (3d) but not epenthesis (4c). Thus, the CCV/CVCV alternation in Dida seems to be best analyzed as a case of underlying /CCV/ with intrusive vowels. On this account, non-alternating CVCV roots are analyzed as underlyingly /CVCV/ whereas alternating roots are analyzed as /CCV/. The representational difference results in distinct surface patterns, where /CVCV/ roots always surface as [CVCV], while /CCV/ roots surface either as [CCV] or with an optional intrusive vowel as [CVCV].

3 Deletion in Guébie

3.1 Language background

Guébie is an endangered Kru language spoken in the Gagnoa prefecture of southwest Côte d'Ivoire by about 7000 people. There is a related variety spoken across the border in Lakota that is also often called Guébie; however, the variety described here is the Gagnoa-Guébie variety, and more specifically, the Guébie of the village of Gnagbodougnoa. Guébie is sometimes classified as part of the Bété subgroup because Bété is the dominant indigenous language of the Gagnoa prefecture. However, Guébie is much more closely related to Dida than to Bété, as argued by Sande (2018) and as articulated by Guébie speakers, who report that while they can understand some Bété, Guébie is much more similar to Dida.

The data presented here comes from field work in collaboration with the Guébie community of Gnagbodougnoa, Côte d'Ivoire between 2013–2022, including seven in-situ fieldtrips, two years of remote elicitation, and two years of work with native speakers in the US and Canada. An initial study of the CVCV/CCV alternation in Guébie is available in Sande (2017). All Guébie examples are labeled with a three-letter code that corresponds to the speaker who provided the examples, as well as the date in YYYYMMDD format. These labels correspond to file bundles of audio files and field notes in the Guébie archival collection in the California Language Archive (Bodji & Sande 2014).

The vowel inventory of Guébie is quite similar to that of Dida, and is given in Figure 4.

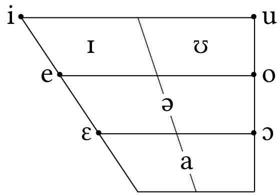


Figure 4: Guébie Vowel inventory

The consonant inventory is provided in Table 3.

Table 3: Guébie consonant inventory

	Bilabial	Lab. dent.	Alveo-pal.	Palatal	Velar	Labialized	Lab. vel.
Plosive	p b		t d	c ɟ	k g	k^{w} g^{w}	kp gb
Nasal	m		n	ɲ	ŋ	ŋ^{w}	ŋm
Fricative		f v	s				
Approx	ɓ		l	j			w

Guébie has four distinct tone heights marked here with numerals 1–4, where 4 is high. Multiple level tones can surface on a single short vowel, resulting in a surface contour tone. Each morpheme is associated with a tone melody underlyingly (except the definite enclitic, whose tone is determined by the tone of the word it attaches to), and the tone melody associates one-to-one with vowels in the word, from left-to-right (Sande 2017). See Sande (2022) for a description of the phonology of Guébie.

Syllables in Guébie typically have the shape CV, though pronouns and loanwords can be vowel-initial, and in certain derived contexts coda nasals appear. Specifically, the vowel of a phrase-final /NV/ sequence can be unpronounced, resulting in a phrase-final nasal coda, in which case the tone of the unpronounced vowel is produced on the preceding vowel. As in Dida, there is a set of words that display CVCV/CCV alternations. Also as in Dida, this alternation is variable and lexically specific.

3.2 Insertion or deletion?

Certain roots can surface as either CVCV or CCV in Guébie. The forms in Table 4 show that while some alternating CVCV sequences have the same vowel and tone on both syllables, and in some cases the second consonant is /l/, there are also alternating forms where the vowels in the two syllables differ (Table 4f,g,h), those where the tones on the two syllables differ (Table 4f,g,h,i), and those where the second consonant is not /l/ (Table 4d,e,g,h,i).[1]

Table 4: CVCV reduced to CCV (syl_20161207)

	CVCV	CCV	
a.	bala$^{3.3}$	bra^{3}	‘hit’
b.	tulu$^{4.4}$	tru^{4}	‘chase’
c.	wʊlʊ$^{3.3}$	wrʊ^{3}	‘granary’
d.	munu$^{3.3}$	mnu^{3}	‘bite/sting’
e.	mana$^{3.3}$	mna^{3}	‘meat’
f.	jɪla$^{2.3}$	jra^{23}	‘ask’
g.	sija$^{2.3}$	sja^{23}	‘be defeated’
h.	kuɓə$^{3.1}$	kɓə^{31}	‘yesterday’
i.	duɓuɓili$^{3.1.1.2.2}$	dɓuɓri$^{3.1.2}$	‘mourning’

[1]There is an alternation between [l] and [r] in Guébie, where [r] is typically used in onset clusters (CCV), and [l] is used elsewhere. While all surface [l] and [r] consonants come from underlying /l/, I use [r] in clusters and [l] elsewhere to reflect production patterns.

Note that in Table 4i, both the first CVCV and the final CVCV sequence can surface as CCV. The second consonant in CCV forms can be /l/ (with the surface form [n] when the preceding consonant is nasal, or [r] immediately after a non-nasal consonant), a glide, or the implosive /ɓ/.

All roots that can surface as CCV have a corresponding CVCV form. Not all CVCV sequences can surface as CCV (Table 5).

Table 5: Non-alternating roots (syl_20161207, syl_20170315)

	CVCV	CCV	
a.	ɟʊla$^{3.2}$	*ɟra^{32}	'take/borrow'
b.	tɛlɪ$^{3.3}$	*trɪ^{3}	'carve'
c.	sijo$^{2.3}$	*sjo^{23}	'wipe'
d.	ɲɛpɛ$^{3.1}$	*ɲpɛ^{31}	'sweep'

Unlike the Dida case, it is not immediately clear given the Guébie facts whether the CVCV/CCV alternation in Guébie is best analyzed as insertion or deletion. V1 of the alternating CVCV sequences need not have the same features as V2 (4f-h). V1 can host its own tone, independent of V2, and in line with regular tonotactics of the language, and [23] contour tones are common in alternating CCV contexts (4g) and non-alternating monosyllabic contexts (/ɲa^{23}/ 'rubber'). V1 in alternating CVCV sequences is not significantly shorter in duration than V1 in non-alternating CVCV sequences.[2] These facts suggest that the initial vowel in alternating CVCV sequences is not intrusive. However, it may still be analyzable as epenthetic (/CCV/ → [CVCV]) or as deleted (/CVCV/ → [CCV]). To determine whether the initial vowels in Guébie alternating CVCV forms can be analyzed as epenthetic, we must determine whether the initial vowels are phonologically predictable.

There is one crucial piece of evidence that leads to the conclusion that no set of phonotactic traits or phonological conditioning factors can predict the initial vowel in alternating CVCV forms. Namely, there are sets of words with the same CCV form, but distinct vowels in their CVCV forms (Table 6).

Given the surface form [jra^{23}], the CVCV form is not predictable; it could be either [jɛla$^{2.3}$] or [jɪla$^{2.3}$] (Table 6). So, in Guébie, the first vowel in a CVCV word is not predictable given the phonological form of its CCV counterpart. This

[2]I do not have access to duration data from Dida to compare with the Guébie duration facts.

Table 6: C1 is not predictable in Guébie (bor_20150602)

	CCV	CVCV	
a.	jra^{23}	jɛla$^{2.3}$	'appear'
b.	jra^{23}	jɪla$^{2.3}$	'ask'

evidence, combined with the fact that the two vowels in alternating CVCV sequences can show independent prosody, lead me to posit that Guébie alternating forms are underlyingly /CVCV/, with optional deletion of the first vowel.

Additional phonological evidence supports the conclusion that the alternating CVCV words in Guébie are not underlyingly /CCV/. The same set of roots that show a CVCV~CCV alternation in Guébie show another phonological alternation: Vowel replacement determined by a subset of suffixes and enclitics. When a third-person object enclitic attaches to an alternating root, both vowels in the alternating root surface with the vowel quality of the enclitic. Alternating roots are shown followed by three distinct third-person object enclitics in Table 7.

Table 7: Object enclitics trigger vowel replacement on alternating roots (syl_20170315)

	Bare verb	3sg.hum =ɔ^{2}	3sg =ɛ^{2}	3pl =ɪ^{2}	
a.	jili$^{2.3}$	jɔl=ɔ$^{2.32}$	jɛl=ɛ$^{2.32}$	jɪl=ɪ$^{2.32}$	'steal'
b.	jɪla$^{2.3}$	jɔl=ɔ$^{23.2}$	jɛl=ɛ$^{23.2}$	jɪl=ɪ$^{23.2}$	'ask'
c.	bala$^{3.3}$	bɔl=ɔ$^{3.2}$	bɛl=ɛ$^{3.2}$	bɪl=ɪ$^{3.2}$	'hit'
d.	wʊla$^{3.1}$	wɔl=ɔ$^{3.12}$	wɛl=ɛ$^{3.12}$	wɪl=ɪ$^{3.12}$	'look at'

The initial vowel in non-alternating CVCV roots does not alternate in this way (Table 8). The final vowel in any CVCV root fails to surface before a V-initial suffix or enclitic due to a regular vowel hiatus resolution strategy, so we are examining alternations in the quality of the initial vowel of the root. Non-alternating roots are shown with the human third-person singular object enclitic in Table 8.

Most third-person pronouns have the shape of a single vowel, but the human 3pl pronoun is /=ʊa^{2}/, produced [=wa^{2}], [ɓatɛ$^{3.1}$] → [ɓat=wa$^{3.12}$], 'search for them'. When this enclitic attaches to an alternating root, the first vowel of the root surfaces as [ʊ] and the second vowel is pronounced [a] (Table 9).

Previous work has analyzed this alternation in third-person object contexts as *vowel harmony* (Sande 2019); however, it may be better analyzed as vowel

Table 8: Non-alternating roots in object contexts (syl_20161207, syl_-20170315)

	Root	Root=ɔ^{2}	
a.	sumu$^{2.2}$	sum=ɔ$^{2.2}$, *sɔmɔ$^{2.2}$	‘boil him’
b.	ɟʊla$^{3.2}$	ɟʊl=ɔ$^{3.2}$, *ɟɔlɔ$^{3.2}$	‘take him’
c.	tɛlɪ$^{3.3}$	tɛl=ɔ$^{3.2}$	‘carve him’
d.	sijo$^{2.3}$	sij=ɔ$^{2.32}$	‘wipe him’
e.	ɲɛpɛ$^{3.1}$	ɲɛp=ɔ$^{3.12}$	‘sweep him’

Table 9: 3pl human object pronouns on alternating roots (syl_20170315, syl_20210817, oli_20210727)

	Root	Root=wa^{2}	
a.	bala$^{3.3}$/	bʊla$^{3.2}$/	‘hit them’
b.	jɪla$^{2.3}$	jʊla$^{2.32}$	‘ask them’
c.	wʊla$^{3.1}$	wʊla$^{3.12}$	‘look at them’

replacement, and not vowel harmony, due to the alternations in the context of alternating roots and third-plural human object pronouns as in Table 9. We do not see the same features on both vowels in Table 9, as would be expected if this were vowel harmony. Rather, we see the two underlying vowels of the pronoun surfacing sequentially in the two available vowel slots.

This phonological evidence suggests that there is a vowel slot in alternating CVCV words, which can be filled in with the first vowel of the /=ʊa/ 3pl.acc object enclitic in vowel replacement contexts, much like Semitic non-concatenative morphology. If alternating words were underlyingly /CCV/, there would be no vowel slot for the first vowel of the 3pl.acc marker to associate to, and we would expect a surface form such as *[bʊlʊa$^{3.3.2}$] (with harmony) or [balʊa$^{3.3.2}$] (without harmony) rather than [bʊla$^{3.2}$] in Table 9a. However, if all alternating morphemes are underlyingly /CVCV/, the vowel replacement facts can be straightforwardly accounted for as non-concatenative association of vowels or vocalic features to a CVCV template.

3.3 Distinguishing alternating from non-alternating CVCV roots

In an analysis of Guébie where alternating roots are underlyingly /CVCV/ with optional V-deletion, the question becomes how to differentiate alternating /CVCV/ roots from non-alternating /CVCV/ roots. Are alternating roots phonotactically distinct from alternating ones? Or is the availability of alternation lexically specific? First, I show that no set of phonotactic features picks out all and only the alternating set of roots.

There are a number of phonotactic traits that alternating roots tend to display, specifically those listed in (6).

(6) Common phonotactic traits of CVCV sequences
- C2 is a sonorant (/l, ɓ, w, j/ (or a nasal in nasal roots))
- V1 and V2 are identical
- T1 and T2 are identical

However, not every root with these features alternates, and not every alternating root has (some subset of) these features. V1 and V2 have identical vowel quality in 329 of the 617 distinct alternating CVCV roots in a Guébie corpus of over 12,000 utterances. We find the same prosody (identical tone) on both syllables in 270 of the 617 alternating forms. If either of these phonotactic traits were diagnostic of CVCV alternation with CCV, we would expect all, or nearly all, alternating roots to show these properties. What we find, though, is that only about half of alternating roots have the same vowel on both syllables in their CVCV form, and fewer than half of alternating roots have the same tone on both syllables.

Perhaps the most crucial piece of evidence that this alternation is lexically specific, and not determined by the phonotactics of a given CVCV form, is that there are (near) minimal pairs of CVCV forms where one alternates and the other does not (Table 10).

The existence of minimal pairs means that at least some information about the availability of /CVCV/ alternation with [CCV] must be lexically specified. No combination of phonotactic traits exclusively and exhaustively predicts whether a root falls into the alternating or non-alternating class. Thus, the Guébie CVCV/CCV alternation is best analyzed as involving underlying /CVCV/ forms that optionally surface as [CCV], where alternating and non-alternating forms are distinguished from each other not based on phonotactics, but lexically.

Any analysis of this deletion phenomenon must account for the fact that this alternation is lexically specific and optional. Deletion is lexically specific in that it does not apply across the board to all lexical items equally, but only applies

Table 10: (Near) minimal pairs of alternating and non-alternating roots (syl_20161207)

	CVCV	CCV	
a.	jili$^{2.2}$	jri^{2}	'be fat'
b.	jili$^{2.2}$	*jri^{2}	'fish'
c.	gɔlɔ$^{3.3}$	grɔ^{3}	'pain'
d.	gɔlɔ$^{2.3}$	*grɔ^{23}	'canoe'
e.	kpolo$^{3.1}$	kpro^{31}	'be clean'
f.	kpoke$^{2.4}$	*kpke^{24}	'crocodile'
g.	ɟulu$^{3.3}$	ɟru^{3}	'salt'
h.	ɟʊla$^{3.2}$	*ɟra^{32}	'take/borrow'

to a subset of lexically specified CVCV sequences. Deletion of the vowel applies optionally in that in any given morphosyntactic and phonological environment, an alternating sequence can surface as either CVCV or CCV.

There are two major classes of analyses that could equally well differentiate alternating from non-alternating CVCV roots in Guébie. One of these types of analysis involves a different underlying representation for alternating and non-alternating forms. One such analysis could involve gradient symbolic representations, where the initial vowel in alternating roots is only partially activated, whereas the initial vowel in non-alternating roots is fully activated (Smolensky & Goldrick 2016). Another such analysis may stipulate that the vowel in alternating roots is represented as defective, lacking the same amount of prosodic structure as the vowels in non-alternating roots (cf. Zimmermann 2013, 2016). The second plausible type of analysis for modeling lexically specific CVCV~CCV alternations involves multiple lexically sensitive phonological grammars or cophonologies (Orgun 1996, Inkelas et al. 1997, Anttila 2002, Inkelas & Zoll 2005, 2007). Cophonology Theory assumes that different morphosyntactic constructions can be associated with distinct phonological constraint rankings or weightings. In Guébie, the class of alternating roots could be associated with a different constraint ranking than the class of non-alternating roots; see Sande (2019) for such an analysis of Guébie vowel alternations. The variability in whether a root surfaces as CVCV or CCV in a given utterance could be modeled in a constraint-based account using Stochastic OT (Boersma 1998, Boersma & Hayes 2001) or Maximum Entropy Harmonic Grammar (Goldwater & Johnson 2003), or through a gestural model (Hall 2003, 2006).

Interestingly, [CCV] forms in Guébie are more common in fast casual speech than in careful speech. Hall's diagnostics (2006) suggest that if speech rate plays a role, vowel intrusion is more likely than vowel epenthesis, but these diagnostics do not make predictions about speech rate and vowel *deletion*. In Guébie, we have seen that the diagnostics for vowel intrusion are not met, and that due to the unpredictabilty of V1, Guébie CVCV/CCV alternations are likely synchronically best analyzed as deletion. Guébie, then, presents a case of speech rate partially determining the output form in a vowel deletion rather than vowel intrusion process, contributing to our understanding of the typological characteristics associated with vowel deletion.

4 Areal findings

The two Kru languages discussed in detail in Sections 2 and 3 were chosen as representative Kru case studies because there is available data on CVCV/CCV alternations, the lexicon, and the phonology of each, and because they serve as representative examples of two common patterns of CVCV/CCV alternation found across Kru languages. This section discusses the synchronic CVCV/CCV alternations across Kru languages, as well as in nearby Mande and Kwa languages, which have been discussed in the literature for decades (cf. Le Saout 1974), providing a picture of the synchronic patterns of CVCV/CCV alternations in West Africa, and drawing conclusions about the diachronic development of these patterns. These areal and historical findings bear on our understanding of the broader typology and development of vowel/∅ alternations across languages (cf. Blevins & Pawley 2010: 40, who present a typology of pathways to vowel loss and vowel insertion, but who do not consider tonal languages or West African languages). Any conclusions made here about proto-forms in Kru, Mande, and Kwa languages should be taken as tentative hypotheses that can be tested in the future with additional careful studies of individual languages and sub-groups like those for Dida and Guébie presented in Sections 2 and 3.

4.1 Kru

In Table 11, a number of Kru languages are listed, along with their classification as Eastern or Western Kru. Each language listed is specified as showing a synchronic CVCV/CCV alternation or not. If there is such an alternation, they are marked as having a phonologically predictable initial vowel quality or not (taken as a diagnostic of whether the alternation involves vowel insertion or deletion), and the possible second position consonants (C2) are provided.

Table 11: CVCV~CCV across Kru

Language	CVCV ~ CCV?	V1 predictable?	Possible C2s
Dida/Vata (Eastern)	Yes	Yes	/l/
Nyabwa (Western)	Yes	Yes	/l, ɓ, w/
Neyo/Neouolé (Eastern)	Yes	Yes	/l/
Guébie (Eastern)	Yes	–	/l, ɓ, j, w/
Grebo (Western)	Yes	–	/l/
Bété (Eastern)	Yes	–	/l/
Godié (Eastern)	Yes	–	/l/
Déwoin (isolate)	–	–	
Kuwaa (isolate)	–	–	

The sources of the generalization, or the data leading to the generalizations in Table 11 include Delafosse (1904), Thomann (1905), Innes (1966), Marchese (1983), Zogbo (1981), Kaye (1981, 1982), Masson (1992), Guéhoun (1993), Egner (1989), Saunders (2009), Yannick (2017), Sande (2017).

Interestingly, the Eastern/Western Kru split does not seem to correspond with a split in behavior of CVCV/CCV forms. There are Eastern and Western Kru languages with predictable V1s in CVCV alternations, which could be analyzed as underlying /CCV/ forms with vowel epenthesis or intrusion, as in Dida. Additionally, there are both Eastern and Western Kru languages with CVCV/CCV alternations but where the quality of the first vowel is not phonologically predictable, as with Guébie. The latter group must be analyzed as having underlying /CVCV/ forms, with optional deletion of the first vowel.

From a diachronic perspective, the data point to synchronically alternating forms across Kru developing from Proto-Kru /CVCV/ forms that alternated with [CCV]. This was reinterpreted as insertion (underlying /CCV/ with predictable [CVCV] variants) in some languages, likely independently in each of the languages where it occurred (Dida, Nyabwa, Neyo). A single, systematic change can result in proto-CVCV forms being reanalyzed as CCV (vowel syncope or deletion), but not the other way around, since the quality of the initial vowel in alternating CVCV/CCV forms is unpredictable in Guébie, Grebo, Bété, and Godié.

Kru isolates like Kuwaa and Déwoin[3], do not show any CVCV~CCV alternation, but have non-alternating CVCV forms that correspond with alternating

[3]Déwoin is sometimes classified as Western Kru, and sometimes as a Kru isolate (Marchese 1983, Zogbo 2012).

forms in Eastern and Western Kru, suggesting that Kuwaa and Déwoin may have split off from the rest of Kru before the CVCV~CCV alternation arose. This alternation is thus proposed to have existed in Kru since before the Eastern/Western split, but it must have arisen after isolates like Kuwaa and Déwoin diverged from the rest of Kru languages.

4.2 Mande

Mande languages are spoken across West Africa, including in Liberia and Côte d'Ivoire where they are in contact with Kru languages. The major sub-division within Mande languages is between Eastern and Western Mande. Many Mande languages, like Kru languages, show synchronic alternations between CVCV and CCV surface forms. In Table 12, Mande languages are listed, along with their classification as (South)eastern or Western Mande. Each language is marked as showing a synchronic CVCV/CCV alternation or not. For languages with synchronic CVCV alternations, I have marked whether the quality of the initial vowel in the CVCV form is predictable given the phonological context, and listed the possible consonants in C2 position. The second consonant, while underlyingly typically only including liquids, can surface as [l], [r], or [n] in most Mande languages, depending on the quality of the first consonant (Vydrine 2004). The sources of the generalizations in Table 12 include Le Saout (1974), Morse (1976), Diallo (1988), Grossmann (1992), Dumestre (2003), Vydrine (2004), Babaev (2011), Sadler (2006 [1949]), Khachaturyan (2015), Vydrina (2017), McPherson (2020).

For Southeastern Mande languages, Vydrine (2004) analyzes the CVCV/CCV alternations as involving underlying /CLV/ forms, which can be realized as CvLV. Realization as CLV or CvLV is variable and not dependent on phonological context. Following a long descriptive tradition, he writes the initial vowel with a lowercase "v" because it is typically pronounced as shorter than other vowels. Similarly, Bearth (1971: 54–56) says the V1 of CVCV forms in Toura (Eastern Mande) is very short and always identical to V2. The first vowel in alternating CVCV forms in Southeastern Mande always matches the features of the second, or is otherwise predictable given phonological context. In Gouro, when the first consonant is velar or palatal, the initial vowel predictably surfaces as round; otherwise, the initial vowel matches the features of the second vowel. From the available examples, it seems that the tone is always identical on both syllables in alternating CVCV sequences. Given all available evidence, it seems quite straightforward to analyze the alternating forms in Southeastern Mande as underlyingly /CCV/, with the option of a phonologically predictable inserted vowel.[4] Given the fact

[4]See also Le Saout (1979) and Vydrine (2010) on *le syllabème* in Mande.

Table 12: CVCV~CCV across Mande

Language	CVCV ~ CCV?	V1 predictable?	Possible C2s
Toura (Southeastern)	Yes	Yes, identical	/l/
Gban (Southeastern)	Yes	Yes, identical	/l/
Dan (Southeastern)	Yes	Yes, identical	/l/
Mano (Southeastern)	Yes	Yes, identical	/l/
Gouro (Southeastern)	Yes	Yes, predictable	/l/
Seenku (Western)	Yes	Yes, schwa	/m,n,ŋ,l/
Dioula (Western)	Yes	–	/l,r/
Zialo (Western)	Yes	–	/l, j, w/
Jalkunan (Western)	Yes	–	/l/
Bambara (Western)	Yes	–	/l,r,n/
Kakabe (Western)	–	–	–
Lɔɔma (Western)	–	–	–
Bandi (Western)	–	–	–
Bobo (Western)	–	–	–

that the initial vowel is always predictable, its presence is optional rather than phonologically determined, it is short or gradiently realized, and it often lacks its own prosody, the best analysis of the inserted vowels in Southeast Mande is that they are intrusive rather than epenthetic (cf. the diagnostics in (3) and (4)).

In some Western Mande languages, some morphemes analyzed as disyllabic, often with a single tone level, act phonologically similar to the CLV morphemes of Southeasterm Mande (Dumestre 2003, Vydrine 2004, Diakite 2006, Green 2010, 2015, 2018). In Bambara, /CVLV/ can be pronounced as [CLV], which, according to Green (2010, 2018), can be analyzed as due to a drive for monosyllabic words (minimality) (recall the discussion in §1.1 of previous work on vowel/∅ alternations analyzed as syncope or deletion motivated by stress and prosody). On Green's analysis, underlying /CVCV/ forms surface as [CCV] when phonotactics allow. Reduction of CVCV to CCV is especially common when the initial vowel of the CVCV sequence is high (/i,u/) (Table 13a-f,j), or if two consecutive vowels have the same vowel quality (Table 13g-i,k). Tones on the two syllables do not need to be identical (Table 13d,j,k). In trisyllabic words, vowel syncope in Bambara can lead to surface [CCV] (Table 13a-c,g-k) or [CVC] syllables (Table 13d-i), depending on the phonotactics (Green 2010: 57–62).

Table 13: Bambara vowel syncope

	CVCV	CCV/CVC	
a.	ká.bi.la	ká.blá	'tribute'
b.	sà.fi.nɛ	sà.fnɛ́	'soap'
c.	sȁ.ku.ra	sȁ.krá	'New Year'
d.	dù.lo.ki	dlò.kí	'shirt'
e.	mò.ri.ba	mòr.bá	'man's name'
f.	sá.nu.ma	sán.má	'holy'
g.	sá.ra.ma	sár.má/srá.má	'famous'
h.	sú.ru.ku	súr.kú/srú.kú	'hyena'
i.	kè.le.ku	kèl.kú/klè.kú	'to stumble'
j.	sí.rã	srã́	'to scar'
k.	dɔ̀.lɔ	dlɔ̌	'beer'

The initial vowel in Bambara alternating CVCV forms is not predictable given the phonological context; it could be a high vowel or a vowel of identical quality to the following vowel. Similarly, the tone of the initial vowel in CVCV forms is not predictable given the CCV form. Thus, in Bambara, /CVCV/ forms can be analyzed as underlying, with the option of surfacing as [CCV]. Other Western Mande languages can be analyzed in the same way as Bambara, with underlying /CVCV/ words that can surface as [CCV] depending on the phonotactic result (Zialo, Jalkunan). Seenku is an interesting case because it seems to be in an intermediate stage between /CVCV/ and /CCV/, where formerly /CVCV/ forms underwent vowel reduction and reanalysis to /CəCV/, and are perhaps in route to being reanalyzed as /CCV/ when the second consonant is a sonorant (McPherson 2020). In other Western Mande languages, surface CVCV forms cannot surface as CCV. This suggests that some Western Mande languages have not yet begun a shift from CVCV to CCV. Vydrine (2004) points out that the Western Mande languages spoken nearby to Southeastern Mande languages, such as Marka/Dafin dialects in Côte d'Ivoire, are more advanced in their shift towards CCV than other Western Mande languages, suggesting an areal or contact effect.

Summarizing the patterns and existing analyses of CVCV/CCV alternations in Mande, Southeastern Mande languages display synchronic CVCV/CCV alternations where the initial vowel of the CVCV form is predictable. They are analyzed as having underlying /CCV/ forms, which are sometimes realized with a short initial vowel [CvCV]. In some Western Mande languages, some underlying /CVCV/ forms can optionally be pronounced [CCV], but there are no underlying /CCV/ forms. In other Western Mande languages there are no surface [CCV] forms at all. These facts suggest that Proto-Mande had CVCV forms, some of which have been reanalyzed as underlyingly /CCV/ in Southeastern Mande, consistent with the conclusions of Vydrine (2004). This reanalysis has not yet occurred in Western Mande, but it seems to be moving in that direction, at least in languages like Seenku, Dioula, Ziallo, Jalkunan, and Bambara.

4.3 Kwa

Some Kwa languages, spoken in the southeast of Côte d'Ivoire just to the east of Kru languages, in southern Ghana, and in central Togo, also show sequences that alternate between CVCV and CCV. Other Kwa languages have surface CCV forms that do not alternate with CVCV. Still others show no surface CCV forms. In Table 14, Kwa language names are listed along with their sub-group. Each language is marked as showing CVCV/CCV alternations or not. If there are CVCV/CCV alternations, Table 14 indicates whether the quality of the first vowel in the CVCV form is phonologically predictable. If there are surface CCV forms at all, the final column indicates which consonants can surface as the second consonant in an alternating CVCV sequence. The second consonant in surface CCV in Kwa is always underlyingly /l/, which can surface as [l], [r], or [n] in many Kwa languages, depending on the phonological environment. The sources of the generalizations in Table 14 come from Westermann (1930), Hyman (1972), Allan (1973), Le Saout (1974), Bergman (1981), Bole-Richard (1983), Dolphyne (1988), Schang (1995), Lenaka (1999), Leben (2002, 2003), Ahua (2004), Bobuafor (2013), Hager-M'Boua (2014), Van Putten (2014), Delalorm (2016), Asante (2017), Paster (2010), Agbetsoamedo (2014), Abunya (2018).

In Selee (Agbetsoamedo 2014), Nkami (Asante 2017), and Twi (Paster 2010) there are stated to be no surface consonant clusters. In Twi, CGV, where G is a glide, is possible on the surface, but is analyzed as underlying CVV (Paster 2010), and does not alternate with CVCV. The other Kwa languages examined here all have surface CCV forms, where underlying /l/ can surface as the second consonant in a surface cluster (Le Saout 1974). When these CCV forms alternate with surface CVCV, the initial vowel is always identical to the second, and is typically very short (Dolphyne 1988, Ahua 2004, Leben 2002). For most Kwa languages in the sample above, there is no attested context where these CCV forms

Table 14: CVCV~CCV across Kwa

Language	CVCV ~ CCV?	V1 predictable?	Possible C2s
Gã (Ga-Dangme)	Yes	Yes, identical	/l/
Baoulé (Potou-Tano)	Yes	Yes, identical	/l/
Agni (Potou-Tano)	Yes	Yes, identical	/l/
Akan (Potou-Tano)	Yes	Yes, identical	/l/
Atchan (Potou-Tano)	(always CCV)	–	/l/
Abidji (Agnegby)	(always CCV)	–	/l/
Avatime (Ka-Togo)	(always CCV)	–	/l/
Tafi (Ka-Togo)	(always CCV)	–	/l/
Leleme (Na-Togo)	(always CCV)	–	/l/
Sekpele (Na-Togo)	(always CCV)	–	/l/
Selee (Na-Togo)	–	–	–
Nkami (Potou-Tano)	–	–	–
Twi (Potou-Tano)	–	–	–

surface as [CVCV]. In Baoulé the alternating words surface as CVCV in slow, careful speech, and in language games, but not in natural, casual speech (Leben 2002, 2003). Even when Baoulé alternating words are pronounced as CVCV in careful speech, the initial vowel is very short, or its duration variable, as expected of intrusive vowels as discussed in §1.1. Similarly, in Gã, Berry & Kotei (1969: 41–42) report that surface CLV syllables, which are quite common, are sometimes produced as CVCV in slow, careful speech, but not in more natural speech.

Across the board in Kwa languages that have surface CCV forms, it is much more common for these words to surface as CCV than as CVCV; whereas in some Kru languages, for example, it is reportedly equally as common for an alternating root to be pronounced as CVCV or CCV in casual speech. For Kwa languages whose CCV forms lack a CVCV alternant, the simplest analysis is that these forms are underlyingly /CCV/, and always surface as such. The question is then how to analyze the alternating forms in languages like Baoulé and Akan, which do show alternations. Due to their behavior in careful speech and language games, Leben (2002, 2003) treats alternating CVCV/CCV words in Baoulé as underlyingly /CVCV/, where the first vowel is reduced or completely elided in most contexts. However, given the fact that the initial vowel is always predictable given the phonological context, and it is always very short unlike other vowels in the language, I propose that it is preferable to analyze CVCV/CCV forms in Baoulé as synchronically underlyingly /CCV/, with intrusive vowels as per Hall (2003, 2006).[5]

In Akan, the identical vowels in alternating CVCV forms need not carry the same tone. When the two vowels of the CVCV form bear distinct tones, the tone typically associated with the first vowel surfaces on C2 in the corresponding CCV form (Dolphyne 1988). For example, the form [ò-fìrí] 'he buys credit' has a low tone on the first syllable of the verb and a high tone on the second. In its CCV form, the low tone surfaces on the /r/, [ò-fr̀í]. Consonants do not typically bear tone in Akan. Additionally, if we said that there is an underlying CCV /fri/ that has a LH tone melody, and that in the case of a contour tone on a short vowel the first tone level is realized on the consonant, we might expect that we would also see CV roots with a tone on the onset consonant. However, no such examples exist in Akan. The simplest explanation of why a tone might surface on the [r] in a CrV syllable, then, is that it is underlyingly a /CVCV/, subject to regular tone association rules followed by optional deletion of the initial vowel segment in casual speech. Additional evidence for underlying /CVCV/ in Akan comes from forms with an initial high +ATR vowel, which show consonant palatalization in their CCV forms: [pira] ~ [p^jra] 'injure' (Dolphyne 1988). There are otherwise no surface palatalized labials in Akan; so, positing that /p^jra/ is underlying would require proposing additional consonants in the Akan inventory that coincidentally only appear in alternating CVCV/CCV forms; whereas, if the CVCV form is underlying, no such requirement is necessary. Thus, based on tonal patterns and the distribution of palatalized labial consonants, it is simpler in Akan to analyze alternating CVCV/CCV forms as underlyingly /CVCV/ with optional deletion. Like in Kru and Mande languages, CVCV/CCV alternations in some Kwa languages are best analyzed as involving underlying /CCV/ plus epenthesis or vowel intrusion (Baoulé), while others are best analyzed as having /CVCV/ forms that undergo optional deletion (Akan).

Of the Kwa languages sampled here, only Gã and Potou-Tano languages, specifically Potou-Tano languages spoken in Côte d'Ivoire, show synchronic CVCV/CCV alternations. (Though not all Potou-Tano languages of Côte d'Ivoire show synchronic CVCV/CCV alternations; see Atchan.) The Potou-Tano languages

[5] One argument Leben makes for /CVCV/ as underlying in Baoulé is that disyllabic verbs show LH tone, which surface CCV verbs also display. CV verbs, on the other hand, have H tone. Thus, in terms of tonal behavior CCV forms are more like CVCV forms than CV forms. This could be due to historical reasons, since CCV forms likely are derived from former CVCV forms, or perhaps the presence of additional segmental structure in CCV forms than CV forms allows them to host an additional tone.

that show synchronic alternations are all in the same sub-group within Potou-Tano, namely, Central Tano, while Atchan and Nkami fall outside the Central Tano group. Nkami and Twi are spoken in Ghana rather than Côte d'Ivoire, so it seems likely that country borders and language contact have played a role in the historical development and stabilization of CVCV/CCV alternations in Kwa. The Potou-Tano languages of Côte d'Ivoire are spoken nearby to Kru and Mande languages which show synchronic CVCV/CCV alternations, supporting an areal story. It seems likely that the CVCV/CCV alternations of non-Kwa languages in Côte d'Ivoire have helped to stabilize the synchronic CVCV/CCV alternations in Ivoirian Potou-Tano languages, whereas CVCV/CCV alternations further away, in Ghanaian Potou-Tano languages, have shifted fully to non-alternating CCV forms. Agnegby and Ka-Togo languages have synchronic CCV forms that do not alternate with CVCV. The Na-Togo languages are mixed between having no surface CCV forms at all on the one hand, and having non-alternating CCV forms on the other.

Hyman (1972: 192) discusses how proto-CVCV forms have become CCV in many Kwa languages over time. He hypothesizes that the existence of nasal vowels in some Kwa languages is due to a further step in the diachronic picture: CVNV words (a subset of alternating CVCV forms) became CNV, which became CṼ. He provides evidence for the intermediate stage from Gwari and Ewe (Hyman 1972: 175–179). In Atchan, there are synchronic nasal vowels and no synchronic CNV syllables (Katherine Russell, p.c.), so it seems to have reached the endpoint of the diachronic chain proposed by Hyman. Hyman's hypothesis raises an interesting question for the Kwa languages that do not have surface CCV sequences at all: Have CVCV words in such languages undergone syncope to CCV followed by consonant fusion or cluster simplification to CV? Or, have CVCV words in such languages never undergone syncope? There seems to be synchronic evidence favoring the former hypothesis. In some Kwa languages, such as Ewe (Westermann 1930), there are both CV and CCV forms of the same word: [ku, klu] 'to scoop', [gbaa, gblaa] 'broad' (p. 20). This variation between CCV and CV suggests that Ewe, and other Kwa languages with no synchronic CVCV/CCV alternation, likely used to have a CVCV/CCV alternation, which then progressed to non-alternating CCV, and then to CCV alternating with CV. Now the proto-CVCV forms in such languages are simply CV: CVCV≫CCV≫CV. For each step in the diachronic process, there is evidence that languages go through a stage of variation. Languages like Akan show synchronic variation between CVCV and CCV, and can be analyzed as currently undergoing the CVCV≫CCV stage, while languages like Ewe show synchronic variation between CCV and CV and can be analyzed as currently undergoing the CCV≫CV stage.

4.4 Areal summary

CVCV/CCV alternations are undoubtedly an areal phenomenon of West Africa, present in at least Mande, Kwa, and Kru languages, especially in the area around Liberia, Côte d'Ivoire, and Burkina Faso. Within these three language families, languages spoken further from the central contact zone in Côte d'Ivoire are more likely to lack a synchronic CVCV/CCV alternation.

In each of the three language families examined here, there are both individual languages that are synchronically best analyzed as having underlying /CVCV/ forms that can optionally delete the first vowel, and there are other languages that are best analyzed as having underlying /CCV/ forms that optionally insert a vowel. In all three language families, as has long been recognized (Hyman 1972, Le Saout 1974, Leben 2002, 2003, Vydrine 2010), there is evidence for proto-CVCV forms, which synchronically correspond to alternating CVCV/CCV forms. These proto-CVCV forms best analyzed as synchronically /CVCV/ with optional vowel deletion in some languages, but have been reanalyzed as underlyingly /CCV/ with optional vowel epenthesis or intrusion in others. Perhaps this areal shift can be explained by a shift towards monosyllabicity, as proposed by Green (2010) for Bambara (Mande). In some Kwa languages, the diachronic shift has progressed even further, with CCV forms undergoing consonant cluster reduction to become CV.

In previous historical work on CVCV/CCV alternations outside of West Africa, Harms (1976) and Hall (2006) show that intrusive vowels have become phonologized over time in languages such as Irish Gaelic, Sardinian, and Finnish: CCV ≫ CvCV ≫ CVCV. The opposite, where phonologically visible vowels are reanalyzed as intrusive over time, seems to have occurred in Dida, but not the closely related Guébie. A similar shift has occurred in certain Mande and Kwa languages. Thus, while there are previous examples of intrusive vowels being phonologized, the Guébie and Dida data show that the opposite is also possible.

Blevins & Pawley (2010) argue that a synchronic CVCV/CCV alternation in Kalam (Trans New Guinea) is the remnant of historical vowel reduction and deletion: (V)CVCV ≫ (V)CvCV ≫ (V)CCV. They summarize the process in this way: "Regular vowel loss yields vowel-zero alternations, which can be reinterpreted as insertions via rule inversion." This cycle seems to be very active in West Africa, where different languages fall synchronically at different states of this diachronic shift. In general, the trend in West Africa seems to be towards monosyllabic words, possibly motivated by a drive towards minimality as suggested by Green (2010, 2018).

5 Conclusions

This paper contributes to the typology of V/∅ alternations by examining two closely related West African tonal languages, and considering both the best synchronic analysis, as well as how they fit into the broader areal and historical picture of V/∅ alternations in West Africa.

I have shown that there is not a single, unified analysis of synchronic CVCV~CCV alternations in Kru languages. In Dida, alternating roots are best analyzed as /CCV/ and subject to vowel intrusion (§2). In Guébie, on the other hand, alternating roots are best analyzed as /CVCV/ (§3), where one class of /CVCV/ roots is lexically specified as alternating (V1 can optionally be deleted), and another class of /CVCV/ forms is not subject to alternation. The alternation cannot go the other direction in Guébie because the initial vowel is not predictable given the phonological environment. These two patterns are shown to be representative of the two most common CVCV/CCV synchronic patterns found across Kru languages (§4.1).

Throughout this paper, evidence for CVCV/CCV alternations as involving deletion or insertion has come from the diagnostics introduced in §1.1.3, summarized in (7). These diagnostics may be useful in distinguishing insertion from deletion in future work on CVCV/CCV alternations in other languages. Note that these are distinct from Hall's diagnostics of whether vowel insertion involves phonologically visible versus invisible vowels (§1.1.1).

(7) Diagnostics for insertion versus deletion in CVCV/CCV alternations

Predictability of V1's quality: If the quality of V1 is predictable given the phonological environment, it can be derived through vowel insertion (cf. Dida). If the vowel quality is not predictable, it must be underlying (cf. Guébie).

Tonal behavior: If the first vowel in an alternating CVCV form can host its own prosody, it is likely an underlying vowel (cf. Akan tonal behavior in CVCV/CCV alternations, discussed in §4.3). On the other hand, if non-regular tonal patterns arise in CVCV forms, it may be because they are derived from underlying CCV forms as in Dida (discussed in §2).

Independence in phonological alternations: If the first vowel in alternating CVCV forms participates independently from the second vowel in regular phonological processes in the language, it is likely an underlying vowel (cf. vowel replacement in Guébie, described in §3).

We have seen that evidence for insertion versus deletion may come from phonological processes seemingly unrelated to vowel presence or absence, such as vowel replacement in Guébie or tonotactics in Dida and Akan. Thus, it is useful to understand the full phonological picture of a language when diagnosing whether a vowel/∅ alternation is best analyzed as deletion versus insertion.

The Guébie data in §3 is dependent on speech rate but was shown to involve vowel deletion rather than insertion. We know from Hall (2003, 2006) that speech rate may play a role in determining whether intrusive vowels are realized or not, but the Guébie facts show that speech rate can also play a role in whether vowels are deleted or not.

From a Kru-internal diachronic perspective, this study has shown that before Eastern and Western Kru split, there was likely a CVCV~CCV alternation best analyzed as deletion. This alternation was reinterpreted as involving underlying /CCV/ forms in some languages (cf. Dida, Nyabwa). Kuwaa and Déwoin, Kru isolates which show no CVCV/CCV alternation, likely split off from the rest of Kru before the CVCV~ CCV alternation arose. In order to confirm this historical hypothesis, additional evidence from other areas of the grammar should be considered.

The Dida case study shows that phonologically visible vowels can be reanalyzed as intrusive over time. This is the opposite of what has previously been shown for a number of non-African languages (Harms 1976, Hall 2006), showing that there is a possible diachronic cycle from /CVCV/ being reanalyzed as /CCV/ and then possibly again being reanalyzed as /CVCV/.

Initial evidence suggests that in nearby Mande languages, proto-/CVCV/ has been reanalyzed as /CCV/ in Southeastern Mande, and is moving towards a reanalysis in many Western Mande languages. One possible explanation is an areal drive towards monosyllabicity (cf. Green 2010, 2015, 2018). In most Kwa languages, also spoken in West Africa, there seems to have been a reanalysis of proto-CVCV forms as CCV, some of which still allow optional epenthesis or intrusion resulting in surface CVCV forms. In some Kwa languages, this change has progressed further, resulting in CCV forms reducing to CV.

Abbreviations

C	Consonant	N	Nasal
V	Vowel	G	Glide
L	Liquid	T	Tone

Acknowledgments

Thanks to Katherine Russell for being an excellent research assistant and compiling comparative Kru data relevant to this project. Thanks to Nancy Hall, Will Leben, and Lynell Zogbo for helpful discussion of related topics. Thanks also to audiences at the LSA 2020 Annual Meeting, WOCAL 9, and the Leipzig Strength in Grammar workshop for comments on related work, and to the audience of the Stonybrook Epenthesis Workshop in September 2021 for their comments on this work. Thanks especially to the Guébie community of Gnagbodgounoa, Côte d'Ivoire for their time and willingness to share the Guébie data presented in this paper. This work was funded by NSF grant #2236768.

References

Abunya, Levina N. 2018. *Aspects of Kaakye grammar*. University of Ghana. (Doctoral dissertation).

Agbetsoamedo, Yvonne. 2014. *Aspects of the grammar and lexicon of Sɛɛɛ*. Department of Linguistics, Stockholm University. (PhD Thesis).

Ahua, Mouchi Blaise. 2004. *Conditions linguistiques pour une orthographe de l'agni: Une analyse contrastive des dialectes sanvi et djuablin*. Universität Osnabrück. (Doctoral dissertation).

Alan, CL. 2000. Stress assignment in Tohono O'odham. *Phonology* 17(1). 117–135.

Allan, Edward Jay. 1973. *A grammar of Buem: The Lelemi language*. University of London, School of Oriental & African Studies. (Doctoral dissertation).

Anderson, James M. 1965. A study of syncope in Vulgar Latin. *Word* 21(1). 70–85.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory* 20(1). 1–42.

Asante, Rogers Krobea. 2017. Introducing Nkami: A forgotten Guang language and people of Ghana. *Legon Journal of the Humanities* 28(2). 73–104.

Babaev, Kiril Vladimirovič. 2011. *Âzyk zialo: Očerk gramatiki i slovar'*. Rossijskij gosudarstvennyj gumanitarnyj universitet. (Doctoral dissertation).

Bearth, Thomas. 1971. *L'énoncé toura (Côte d'Ivoire)*. Summer Institute of Linguistics.

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Bergman, Richard. 1981. *An outline of Igede grammar*. Dallas: SIL.
Berry, Jack & Nii Amon Kotei. 1969. *An introductory course in Ga*. US Office of Education, Institute of International Studies.
Blevins, Juliette & Andrew Pawley. 2010. Typological implications of Kalam predictable vowels. *Phonology* 27(1). 1–44.
Bobuafor, Mercy. 2013. *A grammar of Tafi*. Leiden University. (Doctoral dissertation).
Bodji, Sylvain & Hannah Sande. 2014. *Guébie fieldwork collection*. 2014–15. University of California, Berkeley. (Survey of California and Other Indian Languages). DOI: 10.7297/X208639V.
Boersma, Paul. 1998. *Functional phonology*. Netherlands Graduate School of Linguistics. (Doctoral dissertation).
Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1). 45–86.
Bole-Richard, Rémy. 1983. Ebrié. *Atlas des langues kwa de Côte d'Ivoire* 1. 307–357.
Clements, Nick. 1986. Syllabification and epenthesis in the Barra dialect of Gaelic. In Koen Bogers, Harry van der Hulst & Marten Mous (eds.), *The phonological representation of suprasegmentals*, 317–336. Dordrecht: Foris publications.
Clements, Nick. 1991. Place of articulation in consonants and vowels: A unified theory. 5. 677–123.
Delafosse, Maurice. 1904. *Vocabulaires comparatifs de plus de 60 langues ou dialectes parlés à la Côte d'Ivoire*. Paris: E. Leroux.
Delalorm, Cephas. 2016. *Documentation and description of Sekpele: A Ghana-Togo mountain language of Ghana*. SOAS University of London. (Doctoral dissertation).
Diakite, Boubacar. 2006. *The synchronic link between onset clusters and codas in Bambara*.
Diallo, Mamadou. 1988. *Éléments de systématique et de dialectologie du Marka-kan (Burkina-Faso)*. Université Stendhal Grenoble III. (PhD Thesis).
Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra: Ghana Universities Press.
Dumestre, Gerard. 2003. *Grammaire fondamentale du Bambara*. Paris: Karthala.
Eberhard, David M., Gary F. Simons & Charles D. Fenning (eds.). 2020. *Ethnologue: Languages of the world*. 23rd edn. Dallas: SIL International.
Egner, Inge. 1989. *Précis de grammaire Wobé* (Tome XV). Abidjan: Annales de l'Université d'Abidjan, H, Linguistique.
Fitzgerald, Colleen M. 1997. *O'odham rhythms*. The University of Arizona. (Doctoral dissertation).

Fitzgerald, Colleen M. 2002. Tohono O'odham stress in a single ranking. *Phonology* 19(2). 253–271.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm Workshop on Variation Within Optimality Theory*, 111–120.

Green, Christopher R. 2010. *Prosodic phonology in Bamana (Bambara): Syllable complexity, metrical structure, and tone*. Indiana University. (Doctoral dissertation).

Green, Christopher R. 2015. The foot domain in Bambara. *Language* 91(1). e1–e26.

Green, Christopher R. 2018. *Headedness and prosodic restructuring in Mande*. Presentation at the 49th Annual Conference on African Linguistics.

Grossmann, Rebecca Suzanne. 1992. *Discourse grammar of Bandi*. (Doctoral dissertation).

Guéhoun, N. Augustin. 1993. *Description systématique du Dida de Lakota (langue kru de Côte d'Ivoire)*. Université Stendhal (Grenoble 3). (Doctoral dissertation).

Hager-M'Boua, Ayé Clarisse. 2014. *Structure de la phrase en Abidji, langue Kwa de Côte d'Ivoire*. Université de Genève. (Doctoral dissertation).

Hale, Kenneth & Josie White Eagle. 1980. A preliminary metrical account of Winnebago accent. *International Journal of American Linguistics* 46(2). 117–132.

Hall, Nancy. 2003. *Gestures and segments: Vowel intrusion as overlap*. University of Massachusetts Amherst. (Doctoral dissertation).

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Halle, Morris. 2000. Distributed morphology: Impoverishment and fission. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.), *Research in Afroasiatic grammar*, 125–150. Amsterdam: Amsterdam; Philadelphia; J. Benjamins Pub. Co; 1999.

Harms, Robert T. 1976. The segmentalization of Finnish 'nonrules'. *Texas Linguistic Forum* 5. 73–88.

Hawkins, W Neil. 1950. Patterns of vowel loss in Macushi (Carib). *International Journal of American Linguistics* 16(2). 87–90.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies.* Chicago: University of Chicago Press.

Hérault, G. 1971. *L'aïzi: Esquisse phonologique et enquête lexicale* (Documents Linguistiques 20). Abidjan: Institut de Linguistique Appliquée, Unversité d'Abidjan.

Hyman, Larry M. 1972. Nasals and nasalization in Kwa. *Studies in African linguistics* 3(2). 167–205.

Inkelas, Sharon, Orhan Orgun & Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In J. J. McCarthy (ed.), *Optimality theory in phonology: A reader*, 542–551. Oxford: Blackwell.

Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: Doubling in morphology*, vol. 106. Cambridge: Cambridge University Press.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45.1. 133–171.

Innes, Gordon. 1966. *An introduction to Grebo.* London: School of Oriental & African Studies, University of London.

Jacobs, Haike. 2004. Rhythmic vowel deletion in OT: Syncope in Latin. *Probus* 16. 63–90.

Kager, René. 1997. Rhythmic vowel deletion in Optimality Theory. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 463–499. Oxford: Oxford University Press.

Kawahara, Shigeto. 2007. Copying and spreading in phonological theory: Evidence from echo epenthesis. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.), *University of Massachusetts occasional papers in linguistics 32: Papers in Optimality Theory III*, 1–33. Amherst: GLSA Publications.

Kaye, Jonathan. 1982. Les dialectes dida. In Jonathan Kaye, Hilda Koopman & Dominique Sportiche (eds.), *Projet sur les langues Kru*, 233–295. Montreal: Quebec University.

Kaye, Jonathan & Monik Charette. 1981. Tone sensitive rules in Dida. *Studies in African Linguistics* 8. 82–85.

Kaye, Jonathan D. 1981. La sélection des formes pronominales en vata. *Revue québecoise de linguistique* 11. 117–135.

Khachaturyan, Maria. 2015. *Grammaire du mano.* UC Berkeley. (Doctoral dissertation).

Le Saout, Joseph. 1974. Remarques phonético-phonologiques sur les séquences de type CvLV. 7(1). 5–20.

Le Saout, Joseph. 1979. *Notes sur la phonologie du Gouro (zone de Zuénoula).* Nice: CEPLAN.

Leben, William R. 2002. Tonal feet. In Ulrike Gut & Dayfdd Gibbon (eds.), *Typology of African prosodic systems*, vol. 1, 27–40. Bielefeld: Universität Bielefeld.

Leben, William R. 2003. Tonal feet as tonal domains. In John Mugane (ed.), *Trends in African linguistics 5: Linguistic typology and representation of African languages*, 129–138. Trenton: Africa World Press.

Lenaka, Ngoran Loveline. 1999. *A sketch outline of the phonology of Ndemli*. Universite de Yaoundé Yaoundé. (Doctoral dissertation).

Marchese, Lynell. 1983. *Atlas linguistique Kru*. Abidjan: ILA.

Marchese, Lynell. 1986. *Tense/aspect and the development of auxiliaries in Kru languages*. Dallas: Summer Institute of Linguistics.

Masson, Denis. 1992. Esquisse phonologique du dida de Yocoboué. In Ingeborg Egner (ed.), *Esquisses phonologiques de trois langues ivoriennes: Beng, Dida, Yaouré*, 1–56. Abidjan: Inst. de Linguistique Appliquée (ILA).

McCarthy, John J. 2008. The serial interaction of stress and syncope. *Natural Language & Linguistic Theory* 26(3). 499–546.

McPherson, Laura. 2020. *A grammar of Seenku*. Berlin: De Gruyter Mouton.

Miner, Kenneth L. 1979. Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International journal of American linguistics* 45(1). 25–33.

Miner, Kenneth L. 1989. Winnebago accent: The rest of the data. *Anthropological Linguistics* 31. 148–172.

Morse, Mary Lynn Alice. 1976. *A sketch of the phonology and morphology of Bobo (Upper Volta)*. New York: Columbia University.

Al-Mozainy, Hamza Q̣ublan. 1981. *Vowel alternations in a Bedouin Hijazi Arabic dialect: Abstractness and stress*. The University of Texas at Austin. (Doctoral dissertation).

Orgun, Cemil. 1996. *Sign-based morphology and phonology with special attention to Optimality Theory*. UC Berkeley. (Unpublished dissertation).

Paster, Mary. 2010. The verbal morphology and phonology of Asante Twi. *Studies in African Linguistics* 39(1). 78–99.

Riggle, Jason. 2006. Infixing reduplication in Pima and its theoretical consequences. *Natural Language & Linguistic Theory* 24(3). 857–891.

Sadler, Wendy. 2006 [1949]. A complete analysis of the Lɔɔma language, interior Liberia, West Africa. *Mandekan* 42.

Sande, Hannah. 2017. *Distributing morphologically conditioned phonology: Three case studies from Guébie*. UC Berkeley. (Doctoral dissertation).

Sande, Hannah. 2018. Classification of Guébie within Kru. In Jason Kandybowicz, Travis Majorand, Harold Torrence & Philip T. Duncan (eds.), *African linguistics on the prairie: Selected papers from the 45th Annual Conference on African Linguistics*, 37–49. Berlin: Language Science Press.

Sande, Hannah. 2019. A unified account of conditioned phonological alternations: Evidence from Guébie. *Language* 95(3). 456–497.

Sande, Hannah. 2022. The phonology of Guébie. *Language and Linguistics Compass.*

Saunders, Philip Alexander. 2009. *Encountering Kouya: A study of the key phonological syntactic and discourse properties of an endangered Kru language.* University of Ulster. (Doctoral dissertation).

Schang, Emmanuael. 1995. *Description d'une langue de Côte d'Ivoire: L'avikam.* U.F.R. de sciences du langue, Université de Nancy II. (Doctoral dissertation).

Smolensky, Paul & Matthew Goldrick. 2016. *Gradient symbolic representations in grammar: The case of French liaison.* Ms. Johns Hopkins University and Northwestern University. Available as ROA 1286.

Stanton, Juliet & Sam Zukoff. 2018. Prosodic identity in copy epenthesis. *Natural Language & Linguistic Theory* 36(2). 637–684.

Thomann, G. 1905. *Essai de manuel de la langue néoulé.* Paris: Hachette Livre BNF.

Van Putten, Saskia. 2014. Left dislocation and subordination in Avatime (Kwa). In Rik van Gijn, Saskia Putten, Dejan Matić & Jeremy Hammond (eds.), *Information structure and reference tracking in complex sentences*, 71–98. Amsterdam: John Benjamins.

Vogler, Pierre. 1976. *Description synchronique d'un parler kru: Le vata.* Paris: Sorbonne.

Vogler, Pierre. 2015. Le sèmè/siamou n'est pas une langue kru. *Hal open science, hal-01182225.*

Vydrina, Alexandra. 2017. *A corpus-based description of Kakabe, a Western Mande language: Prosody in grammar.* Sorbonne Paris Cité. (Doctoral dissertation).

Vydrine, Valentin. 2004. Areal and genetic features in West Mande and South Mande phonology: In what sense did Mande languages evolve. *Journal of West African Languages* 30(2). 113–126.

Vydrine, Valentin. 2010. Le pied métrique dans les langues mandé. *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels.* 53–62.

Westermann, Diedrich. 1930. *A study of the Ewe language.* Oxford: Oxford University Press.

Willett, Elizabeth. 1982. Reduplication and accent in Southeastern Tepehuan. *International Journal of American Linguistics* 48(2). 168–184.

Willett, Thomas Leslie. 1991. *A reference grammar of Southeastern Tepehuan.* Dallas: SIL.

Yannick, M. Allou Allou Serge. 2017. *Description linguistique et sociolinguistique de kouzie: Parler kru du Côte d'Ivoire.* Université Felix Houphouët Boigny. (Doctoral dissertation).

Zimmermann, Eva. 2013. Non-concatenative allomorphy is generalized prosodic affixation: The case of Upriver Halkomelem. *Lingua* 134. 1–26.

Zimmermann, Eva. 2016. The power of a single representation: Morphological tone and allomorphy. *Morphology* 26(3-4). 269–294.

Zogbo, Gnoléba Raymond. 1981. *Morpho-syntaxe, lexicologie des bases, lexicologie des nominaux du bété (région de daloa)*. Thèse de 3ème cycle, Sorbonne. plate 7.1 Gur languages. (Doctoral dissertation).

Zogbo, Lynell Marchese. 2012. Kru revisited, Kru revealed. In *Towards proto-Niger-Congo: Comparison and reconstruction*, 18–21. Paris: International Congress.

Zogbo, Lynell Marchese. 2019. Central vowels in the Kru language family: Innovation and areal spreading. In *Theory and description in African Linguistics*, 725–750.

Chapter 3

Vowel predictability and omission in Anindilyakwa

John Mansfield[a], Rosey Billington[b] & Hywel Stoakes[a]

[a]The University of Melbourne [b]The Australian National University

The Australian language Anindilyakwa has some vowels that are to a large extent contextually predictable, and arguably epenthetic. In previous work on the language there are differing views on the segmental contrasts, phonotactic patterns and lexical representations of these vowels. Drawing on information-theoretic approaches, we investigate the predictability of vowel occurrence across different consonant environments in Anindilyakwa words, using orthographic representations from an existing wordlist, and speech production data collected with seven Anindilyakwa speakers. We find that there is a high level of predictability in the word-internal occurrence of non-low vowels compared to low vowels. At the same time, the word-final low vowel *-a* is completely predictable. In our speech production data we find that the predictable vowels (both word-internal non-low vowels, and the final *-a*) are quite frequently omitted even in a relatively careful speech register, while the unpredictable vowels are never omitted. Our findings support previous research that draws a connection between segmental predictability and phonetic reduction or deletion, and we show that this association extends to segments that can be analysed as epenthetic.

1 Introduction

In this chapter we focus on one particular aspect of epenthesis, namely the *predictability* of epenthetic segments. We present new data and analysis on Anindilyakwa, a northern Australian language in which the presence of particular vowel phones is to a large extent contextually predictable. In part this involves canonically epenthetic vowels, inserted predictably at morphological boundaries

John Mansfield, Rosey Billington & Hywel Stoakes. 2024. Vowel predictability and omission in Anindilyakwa. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 57–84. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264532

and the end of words. But vowel predictability in Anindilyakwa also involves morpheme-internal positions, where vowels that can be grouped as "non-low" are to a large extent predictable according to consonantal context. We can thus posit redundancy-free representations in which all boundary vowels and many internal vowels are omitted, as in (1).[1] In this chapter we investigate both types of predictability, following an information-theoretic approach to phonological analysis.

(1) nəŋərəŋkənama
nŋ-rŋk-na-m
1s-see-NPST-MUT
'I see.'

We begin this chapter by briefly surveying the common ground between information theory and phonological analysis, showing that this has particular relevance to epenthesis (Section 2). We then introduce the Anindilyakwa language and its segmental phonology (Section 3), before presenting new data on this language. Our first original contribution investigates the extent to which the appearance of non-low vowels is predictable according to context, based on orthographic forms in a wordlist (Section 4). Our second contribution uses new field recordings to analyse the variable production of vowels in elicited sentences, showing that more predictable segments are more likely to be omitted (Section 5). In the final section we interpret our findings with respect to theories of lexical representation, and point out some directions for further research (Section 6).

2 Phonological predictability, phonetic reduction and epenthesis

The core concerns of phonology have substantial overlap with those of information theory (IT). Both disciplines analyse how sets of distinct symbols are combined into sequences to convey meanings, with concomitant issues such as contrast, redundancy and contextual predictability (Hockett 1967, Goldsmith 2000). One salient difference between standard phonological analysis, and IT analysis, is the use of discrete vs gradient methods. Phonology is mostly concerned with categorical distinctions (if a pair of phones ever contrast in the same environment, then they are distinct phonemes), whereas IT quantifies *degrees* of contrast,

[1] Our morphological glosses follow the Leipzig Glossing Rules, with one language-particular addition: MUT(ual), a marker of 'symmetrical information access from speaker perspective' (Bednall 2021) (see example (1)).

based on the probabilistic distribution of symbols in context. An IT approach is useful for studying Anindilyakwa vowels because they are largely predictable according to context, while nonetheless exhibiting contrasts in a few lexemes. IT provides formulae for mapping out this 'intermediate' territory between the fully predictable and the fully contrastive (Hall 2009, 2013a, Parker 2015). While IT inherently lends itself to analysing gradient phenomena, there has as yet been little discussion of how it relates to gradient models of phonology such as exemplar theory (Bybee 2000, Pierrehumbert 2001). It is beyond the scope of the current study to bridge the gap between IT and gradient theories of phonology, but in our closing discussion we will argue that IT and gradient lexical representations are fundamentally compatible.

Information theory treats communication as the transmission of messages from a sender to a receiver via symbolic sequences (Hartley 1928, Shannon 1948). The messages could for example be words selected from a lexicon, where each word is uniquely associated with a sequence of phonological symbols. A fundamental constraint of communication is that the symbolic sequences must be relatively short to be usable, and IT captures this constraint by calculating the shortest possible encoding scheme for an array of messages. The calculation of these efficient encodings leads to more general formulae for the information-content of symbols, and the effects of context in symbolic sequences.[2]

To some extent, IT provides an alternative theoretical path to the same types of conclusions as standard phonology. If we have a lexicon where many words contain a substring [tək], but there are no words containing the substring [tk], the intervening schwa does not distinguish between any possible messages. It thus conveys zero information, and can be omitted from an efficient encoding. Phonological analysis might similarly conclude that there is a redundancy-free underlying sequence /tk/, and the schwa is inserted at some level of derivation.

But IT additionally quantifies the amount of information in phonological symbols that are only partially predictable. Given a lexicon where we frequently encounter [tək], but there are also a few words containing [tk], the intervening schwa now carries some quantity of information, because it contributes to distinguishing one word from another. This degree of predictability is formulated as 'surprisal', that is *negative log probability*, where a low surprisal value captures the highly predictable nature of schwa insertion. Surprisal can also be thought of as a quantity of 'information', because unsurprising segments have little capacity to distinguish between messages, and are therefore relatively uninformative (Shannon 1951). Using a binary logarithm, quantities of information are often

[2]For mathematical details, see Shannon (1948) and Cover & Thomas (2002).

expressed as 'binary digits' or 'bits', reflecting the fact that the surprisal of a message is equal to its binary character length in an optimally efficient encoding (Shannon 1948).

IT facilitates various predictions about natural language phonology and phonetics. Among other things, it predicts that phones or phone-sequences should be hyper-articulated when they contain a greater quantity of information (i.e. more discriminative of distinct messages), but hypo-articulated when they contain less information (Lindblom 1990). In recent years, a substantial body of work has shown that when a phone is less informative, its phonetic cues are indeed more likely to be reduced or deleted (e.g. van Son & van Santen 2005, Hall 2009, Cohen Priva 2015, 2017, Shaw & Kawahara 2017, Hall et al. 2018).

Epenthesis is typically defined as a relationship between underlying lexical representations and surface phonology, whereby a segment is present on the surface but not in the underlying representation. While much of the research on epenthesis focuses on its function to satisfy phonological constraints (Hall 2011), in this chapter we focus on its predictable character. In general, segments are only considered epenthetic if they occur predictably in some context, whereas segments that are distinctive to particular lexical items are considered to be part of the underlying lexical representation. If a purportedly epenthetic vowel ceases to be predictable, for example through sound changes in the conditioning consonants, then its epenthetic status is called into question (Hall 2011: 1579). Since epenthetic segments are by their nature predictable, IT would therefore suggest that epenthetic segments should be more subject to phonetic reduction or omission, compared to more lexically informative phones – and this is exactly what has been found, at least in some cases of epenthesis (Hall 2013b). There has also been some research connecting IT directly with epenthesis, for example arguing that the unusual occurrence in French of [ø] as an epenthetic vowel may be explained by the low informativity of this segment in lexical representations (Hume et al. 2011; see also Tily & Kuperman 2012).

The current study draws on methods developed in IT phonology, applying them to the arguably epenthetic vowels of Anindilyakwa. While previous studies have mostly focused on the contextual predictability between pairs of phones, or the general information content of one particular phone type, we instead measure the predictability of vowel presence vs absence in consonantal contexts. Additionally, we investigate the proposed association of predictability with phonetic reduction and omission, by investigating vowel omission patterns in speech production.

3 The Anindilywaka language and its segmental phonology

Anindilyakwa is a Gunwinyguan language (Glottocode: anin1240), owned and spoken by the Warnumamalya people of the Groote Archipelago in northern Australia. Warnumamalya generously shared information about Anindilyakwa that is the basis for this study, provided recordings of their speech to the first author, and supported this research especially through the Groote Eylandt Language Centre.

There are several previous studies of Anindilyakwa phonology (Heath 2020, Stokes 1981, Leeding 1989, van Egmond 2012, van Egmond & Baker 2020). In this chapter we draw on all these previous works, but perhaps the most important is the earliest, that being Heath's unpublished sketch where he observes that many Anindilyakwa vowels are largely predictable in their occurrence and quality. This is relatively unusual among Australian languages (but see also the proposals of 'vertical vowel systems' in Arandic languages, e.g. Breen & Dobson 2005).

Anindilyakwa has a small phonemic inventory, highly constrained phonotactics, and unusually long words. All words begin with a consonant or a low vowel, and in citation form all words end in *-a*. Word-internally there is a relatively constrained range of consonant clusters (Heath 2020, van Egmond 2012), which partly motivates the analysis of vowel epenthesis to be described below. Words tend to be long (Leeding 1989: 68), even when they denote relatively basic concepts, e.g. *nεɲcarŋaʎiʎa* 'boy', *εŋkəparŋʷarŋʷa* 'heavy' and *al̪uŋkᵚuwaruwaʎa* 'shade' .

There have been various analyses of the underlying vowel phonemes and patterns of surface allophony in Anindilyakwa, with previous work noting vowel frontness and rounding is conditioned by the place-of-articulation of neighbouring consonants. Partly due to the differing analyses, the same lexical items are sometimes represented with different orthographic vowels in different sources. Despite different phonemic analyses, there is broad agreement that (surface) vowel phones in Anindilyakwa include two LOW vowels [ε, a], and three NON-LOW vowels [i, ə, u].

There are also various analyses of the consonant phonemes, but for our study we use the system represented in Table 1. It is broadly typical of Australian languages, with a large number of place contrasts, and more sonorants than obstruents (Butcher 2006). The most recent analysis of the consonant inventory posits complex segments, namely a prenasalised stop series /mp, nt, n̪t̪, ɳʈ, ɲc, ŋk, ŋkʷ/, and labio-velar /kp, ŋm, ŋp/ (van Egmond 2012; van Egmond & Baker 2020), but we treat these instead as clusters.[3] This is firstly because like other clusters they appear only in word-internal positions; secondly because their articulation seems to us much like clusters in other Australian languages; and thirdly because of phonotactic patterns that form the core of the analysis below (see Section 4.1).

Table 1: Anindilyakwa consonant inventory

	Labial	Anterior	Retroflex	Alveo-palatal	Dorsal	Labialised dorsal
Stop	p	t̪ (t)	ʈ	c	k	kʷ
Nasal	m	(n̪) n	ɳ	ɲ	ŋ	ŋʷ
Lateral		l̪ (l)	(ɭ)	ʎ		
Trill		r				
Approx	w		ɻ	j		

As observed by Heath and subsequent analysts, non-low vowels are generally realised as [u] when adjacent to a labial or labialised dorsal consonant, and [i] when adjacent to an alveo-palatal consonant. Both contexts are exemplified in [jipuɻat̪a] 'wallaby'.[4] Elsewhere, non-low vowels are generally realised as [ə], as for example in [t̪ənəna] 'mosquito'. At the same time, various lexical exceptions have been observed (Stokes 1981), especially in instances where [i] has no conditioning palatal, as in [məɻirpa] 'back'.[5] There is also some evidence for conditioning in the low vowels, where [ɛ] frequently appears adjacent to a palatal and is arguably an allophone of [a] in this context, as in [micijɛʎa] 'beach', though this pattern appears to be weaker than the non-low vowel conditioning.

The predictability of non-low vowel quality suggests a phonological analysis where there is a single underlying non-low vowel, with surface allophony accounting for [i ~ ə ~ u] (Leeding 1989). But Heath goes further than this, additionally proposing that since Anindilyakwa has pervasive cluster constraints in its

[3]Also our 'Anterior' place of articulation merges potential distinctions between apical-alveolar and lamino-dental articulations, since we did not observe any contrasts between these in our field data, and in any case such contrasts are considered marginal (for discussion see van Egmond 2012).

[4]There is no consensus on how they are realised when these two conditions overlap, e.g. following a labial and preceding a palatal, and there are differences in the observations of vowel distributions relative to specific consonant segments within broad place categories.

[5]Our production data includes several of these lexemes with 'un-conditioned' close front vowels. Some, such as *ɛniŋapa* 'good', show phonetic variation [ɛniŋapa ~ ɛnəŋapa]. But *məɻirpa* 'back' appears to have a quite consistent close front vowel. Further investigation of vowel quality is an important topic for further research.

surface phonology, the very presence of non-low vowels can be considered epenthetic – a product of surface articulation, but absent in underlying lexical representations such as /jpɻat̪/ → [jipuɻat̪a]. A more canonical type of vowel epenthesis is also observed at word-internal morphological boundaries, where non-low vowels are predictably interpolated as in /k-wʈa-n-m/ → [kuwuʈanəma] 'will climb (IRR-climb-NPST-MUT)'. Heath also reports that the insertion of non-low vowels between consonants is somewhat variable, and their phonetic character is 'brief and indistinct' (Heath 2020, section 1.8). These claims are supported in our speech production study below.

As shown in the examples above, there is another another aspect of vowel predictability in Anindilyakwa: every word in the lexicon ostensibly ends with *-a*. This suggests that while non-low vowels are epenthesised word-internally, the low vowel *-a* is epenthesised word-finally (see van Egmond 2012 for further discussion).

To summarise the existing analyses of Anindilyakwa vowels, the top line in (2) shows the surface form of a noun with a case suffix. The next line down reflects the view shared by all previous work that both the non-low vowel at the morphological boundary, and the low vowel at the end of the word, are epenthetic. The line below that reflects an additional proposal (as in Leeding 1989) that there is a single underlying non-low vowel, with surface allophones conditioned by neighbouring consonants. The line below that reflects Heath's additional proposal that morpheme-internal non-low vowels might be considered epenthetic. Since non-low vowels are uncontroversially epenthesised at word-internal morphological boundaries, the implication is that all (word-internal) non-low vowels may be treated as epenthetic. This proposal is the most efficient from an IT point of view, since it reduces the length of symbolic representations, and it is this proposal that we examine below through the lens of predictability and informativity.

(2) jipuɻat̪əjat̪a (surface)
jipuɻat̪-jat̪ (boundary epenthesis)
jəpəɻat̪-jat̪ (non-low allophony)
jpɻat̪-jat̪ (non-low epenthesis)
wallaby-PURP
'For wallaby.'

While Anindilyakwa non-low vowels can be analysed as synchronically epenthetic, historically they derive from reduction of full vowels *i, *u and *a in coarticulatory and unstressed environments, as shown in recent work reconstructing the ancestor of Anindilyakwa and closely-related Wubuy (van Egmond & Baker

2020: 159). This suggests an alternative synchronic analysis of non-low vowels as being variably deleted, rather than variably inserted. While this analysis is quite plausible, for most of this chapter we focus instead on the epenthetic analysis as this reflects our interest in the efficiency of redundancy-free coding. However, in our closing discussion we will further probe the question of lexical representations, and argue that probabilistic, gradient approaches ultimately avoid having to make an analytic choice between insertion and deletion.

Other elements of Anindilyakwa phonology remain somewhat under-studied, especially prosodic phonology (but see Leeding 1989: 138–141; van Egmond 2012: 27ff.). While we suspect that higher prosodic structures may have an important role in the patterning of non-low vowels, this dimension of the system must await further research.

4 Predictability of vowels in Anindilywaka lexical data

In this section we investigate vowel predictability in the Anindilyakwa lexicon, using a wordlist of 3038 orthographically-represented lexical items drawn from a dictionary (Waddy 1989). From this, we extract 6943 consonant contexts that are the potential environments for non-low vowels to occur.[6] In this section we illustrate wordlist citation forms using italic text as in *jipuɻat̪a* 'wallaby', transformed into IPA to enable cross-reference to the phonemic inventory outlined above. We use angled brackets to indicate redundancy-free phonological representations as in /jpɻat̪/.

Based on previous analyses, we assume that word-internal low vowels [ɛ, a] are encoded in lexical representations, and therefore wherever two consonants occur in sequence *without* a low vowel in between, we have a consonant context that is the possible site of non-low vowel interpolation. We label such contexts 'C_C', and analyse the predictability of non-low vowel interpolation in the 6943 instances of C_C extracted from the wordlist. For example from the wordlist item *jipuɻat̪a* 'wallaby' we extract the contexts j_p and p_ɻ. We also calculate the predictability of low vowels, given a representation in which non-low vowels are omitted, which confirms that low vowels are much less predictable (i.e., more informative) in this representation.

Technically, we could reverse our analysis and assume that non-low vowels are encoded in lexical representations, then calculate the predictability of low

[6] An estimation of the discourse frequency of these lexical items would also be of interest for an IT analysis, but unfortunately data of this type is not currently available, and we therefore treat all lexemes as equiprobable messages in a communication channel.

vowel interpolation in the remaining C_C contexts. However this would be more difficult to justify, given the previous analyses proposing that non-low vowels are epenthesised word-internally, and low vowels word-finally. We return to this issue in Section 4.2 below.

It would also be interesting to investigate the predictability of non-low vowel *quality* in the wordlist data, i.e. the extent to which [i, ə, u] can be predicted by the place of articulation of C_C contexts, as suggested in previous work. However, we are not able to do this because the qualities [i, ə] are not distinguished in the wordlist, which presents both [i] and [ə] using the character <i>. Conversely, we suspect that some entries may 'over-distinguish' vowel types, for example the r_p context is attested with both <i> and <u> as in *-ripijena* 'see dimly' vs *-rupuriŋkina* 'watch over', but we suspect that such words may have a single vowel type that is phonetically intermediate between [i ~ u]. Analysing the predictability of non-low vowel quality will require quantitative analysis of vowel formants in production data, which is beyond the scope of this chapter, but the subject of preliminary analyses currently underway (Billington et al. 2022).

4.1 Predicting non-low vowels by consonantal context

Our analysis of the orthographic wordlist data shows that the presence or absence of a non-low vowel is to a large extent predictable from the neighbouring consonants. To some extent this is implied by earlier observations that Anindilyakwa permits only a limited range of consonant clusters (both within and across syllables), which suggests that for any other C_C context, a non-low vowel is predictably interpolated. However we here add a further observation: not only are there certain C_C contexts where a vowel is predictably present, but there are also C_C contexts where a vowel is predictably *absent*. That is to say, for a context like m_p that can form a cluster, not only do we find many instances of clusters as in *ŋampuwa* 'where to?', but we almost never find interpolated non-low vowels like **mup*. Thus both presence and absence of non-low vowels is generally predictable from C_C context, and this greatly reduces their overall information value.

The simplest types of C_C context are those in which non-low vowels generally do not occur. These are the most frequent cluster types noted in previous work, namely homorganic nasal-stop sequences [mp, nt, ɳʈ, ɲc, ŋk, ŋkʷ], and dorsal-labial sequences of equal or increasing sonority [kp, ŋm, ŋp]. For homorganic nasal-stop sequences, the wordlist attests just a handful of exceptional vowel interpolations such as <i> in the n_t̪ context in *ɛnit̪ira* 'fins'. For dorsal-labial sequences there is also a handful of exceptions such as <u> in the ŋ_p context in *aŋupina* ‘cloud’. Some of these may be explained by borrowing, for example *aŋupina* < Yolŋu *waŋupini* ‘cloud’, while other apparent exceptions can be explained as a constraint against triple-consonant clusters *[ŋkp], for example *ɛŋkuparŋʷarŋʷa* ‘heavy’.

The predictable absence of non-low vowels in contexts such as m_p and k_p is another reason why we treat these as clusters rather than complex segments (as in van Egmond 2012; van Egmond & Baker 2020). By positing bisegmental sequences such as m_p and k_p, in a system of generalised predictability for non-low vowels, we can explain both the prevalence of clusters such as [mp] and [kp], and the scarcity of sequences such as *[mup] and *[kəp]. The scarcity of interpolated sequences would remain unexplained if we instead posit complex segments such as /mp/ and /kp/. If these were phonemic segments, existing alongside simple segments such as /m, k, p/, then we would need an additional explanation as to the scarcity of *[mup], *[kəp] *etc.* in the lexicon.

Consonant contexts that predictably do have an interpolating non-low vowel require more detailed specification. There is a large set comprising any context with equal or increasing sonority (e.g. t̪ət̪, t̪ən, t̪əl̪, nən, nəl̪…), and these account for about half of all the C_C contexts in the dataset (52%, N=3628). There are just a handful of exceptions to the pattern of vowel interpolation in this context, some of which conform to Heath’s (n.d.: 1) observation of coronal stop clusters [ʈc, t̪c], e.g. *-kpət̪ciji* ‘be on an edge’. Others involve a retroflex nasal followed by a palatal glide, e.g. *-ʎaŋkaɳjɛra*, ‘hard’.

There are also several types of decreasing-sonority C_C that are predictably interpolated. These are listed in Table 2 (parenthetic labels like ‘M_T’ will be used to refer back to these classes below).

Finally, there are the other types of decreasing-sonority C_C, in which non-low vowels are unpredictably present or absent (listed with examples in Table 3). Whereas the predictable C_C types described above can be efficiently encoded without non-low vowel symbols, as in /ŋampw/ ⟶ [ŋampuwa] and /jpɻat̪/ ⟶ [jipuɻat̪a], the unpredictable types require non-low vowels to provide complete lexical representations, for example *nəp* in /akʷanɛnəpk/ ‘life’ vs *np* /l̪amanpcn/ ‘be absorbed’.

Figure 1 shows the token counts of C_C non-low vowel interpolation contexts in the wordlist, grouped according to whether they are predictably clustered ‘CC’, predictably interpolated ‘CəC’, or ‘Unpredictable’. The figure shows that the largest group is those that are predictably interpolated (N=4445, 64% of the data), while smaller groups constitute the predictable clusters (N=1267, 18%) and unpredictable contexts (N=1231, 18%).

Table 2: Decreasing-sonority C_C types with predictable vowel interpolation.

Description	Label	Example
Peripheral or palatal nasal followed by heterorganic stop	M_T	*mamuṯakpa* 'tail'
Palatal lateral followed by stop or nasal	ʎ_T	*-ŋkaʎikəna* 'wet'
Anterior lateral followed by non-labial nasal	l_N	*-ḻaḻəna* 'sit'
Trill followed by coronal nasal	r_N	*-kʷuʎarəna* 'shine'
Approximant followed by nasal (except ɻ_m, see below)	Y_N	*awuɲampa* 'anger'
Approximant followed by liquid	Y_L	*-ḻawuḻawɛna* 'be stretched out'
Non-retroflex approximant followed by stop	Y_T	*wijiṯa* 'straight'
Retroflex approximant followed by dorsal stop	ɻ_k	*akʷaɻaɭəkaja* 'vines'

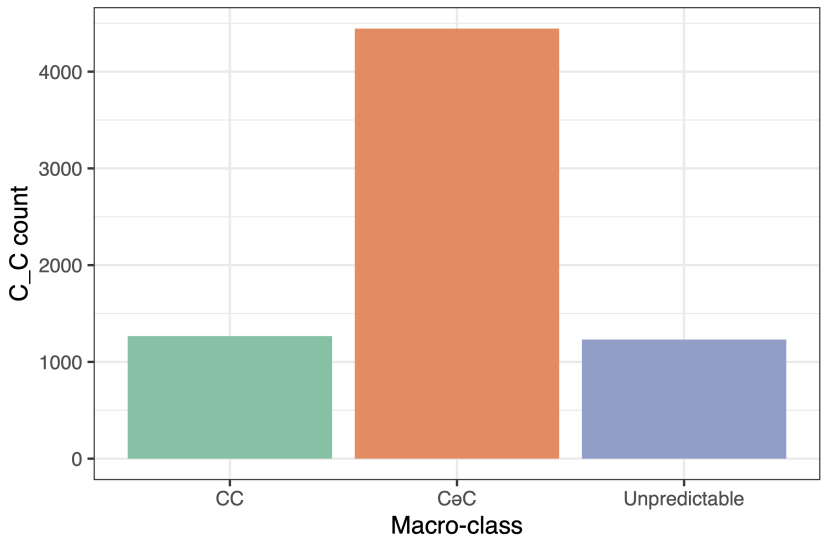


Figure 1: Token counts in the wordlist of C_C contexts grouped according to predictability of non-low vowel interpolation

Table 3: Decreasing-sonority C_C types with unpredictable vowel interpolation.

Description	Label	CƏC example	CC example
Anterior nasal followed by non-heterorganic stop	N_P	*akʷanɛnəpəka* 'life'	*-l̪amanpəcina* 'be absorbed'
Anterior liquid followed by stop	L_T	*-ɛrəpuɻan̪t̪ə* 'not want'	*-ɛrpal̪ici* 'separate'
Anterior liquid followed by bilabial nasal	L_m	*-ŋʷurməl̪ɛna* 'grumble'	*-ŋʷurumɛjina* 'keep quiet'
Trill followed by dorsal nasal	r_ŋ	*-wɛrəŋɛkpuɻakəna* 'comfort'	*-alkarŋi* 'cut grass'
Retroflex approx followed by a non-dorsal stop	ɻ_T	*-ŋʷuciɻət̪əna* 'become deep'	*-aŋmaɻt̪ɛ* 'hate'
Retroflex approx followed by bilabial nasal	ɻ_m	*-apaɻumə* 'search'	*amaɻmara* 'sore'

While Figure 1 already suggests that non-low vowels are to a great extent predictable in the Anindilyakwa lexicon, we can provide an information-theoretic measurement of this predictability. As outlined above, surprisal (negative log probability) can be used as a measure of predictability for a segment in a given context. When we additionally consider the predictability of a choice between phonological possibilities, in this instance clustering versus non-low vowel interpolation, we use the *weighted average surprisal* of the possible outcomes. This quantity is often referred to as "entropy", and can be thought of as the information-value of a paradigmatic choice, such as a phonological contrast. Table 4 illustrates the CəC and CC counts for each context, as well as the entropy value in bits, $H(ə \sim \varnothing)$, quantifying information value of non-low vowel interpola-

Table 4: Predictability of non-low vowels in various C_C types

Context type	CəC count	CC count	$H(ə \sim \varnothing)$
Predictable interpolation			
Equal or increasing sonority	3615	13	0.03
l_N	55	0	0.00
ʎ_T	162	0	0.00
M_T	105	1	0.08
r_N	52	0	0.00
ɻ_k	60	2	0.21
Y_L	133	2	0.11
Y_N	177	0	0.00
Y_T	68	0	0.00
Predictable clustering			
Dorsal_labial	23	392	0.31
Homorganic nasal_stop	16	836	0.13
Unpredictable clustering			
L_m	61	36	0.95
L_T	309	364	0.99
N_P	52	25	0.91
r_ŋ	63	114	0.94
ɻ_m	39	17	0.89
ɻ_T	46	105	0.89
WEIGHTED AVERAGE			0.23

tion in that context.[7] The table also gives the weighted average entropy for C_C contexts in the entire lexicon, where each C_C context is weighted by the probability of this context being encountered. This overall weighted average is 0.23 bits, which is a low figure given that complete predictability would be zero, and the maximum informativity would be 1.

[7]The formula for this entropy, i.e. weighted average surprisal, is:

$$H(ə \sim \varnothing) = (p(ə) \times -\log_2 p(ə)) + (p(\varnothing) \times -\log_2 p(\varnothing)).$$

In fact there is some evidence suggesting that 0.23 bits is actually an overestimate of the informativity of non-low vowels in Anindilyakwa. We will see below (Section 5) that non-low vowels in production data are variably omitted in decreasing-sonority contexts, e.g. for [u] in [mamut̪akpa ~ mamt̪akpa] ‘tail’. This suggests that some of the lexical unpredictability attested in the wordlist for decreasing-sonority contexts may actually reflect phonetic variation, rather than lexical contrasts.

4.2 Predictability of low vowels

To put the information measurement of non-low vowels in perspective, we calculate the amount of information carried by low vowels in the proposed underlying representations. Given that non-low vowels have been shown to be largely predictable, we assume representations in which these are absent from underlying representations, e.g. /jpɻat̪/ ‘wallaby’. This leaves every C_C sequence as a context for potential interpolation of a low vowel, e.g. j_p, p_ɻ, ɻ_t̪. To maximise comparability with the non-low vowels, we set aside the difference between low vowel qualities [a, ɛ], and ask only whether the presence/absence of a low vowel is predictable in various C_C types.

Using the same C_C context types that were identified above, we find that low vowels are much less predictable than the non-low vowels, that is to say they carry much more information. As shown in Table 5, the presence vs absence of a low vowel is relatively unpredictable in almost every C_C context type, including the most frequent contexts such as ‘Equal or decreasing sonority’. The only contexts that have relatively low entropy are dorsal-labial and homorganic nasal-stop sequences, which again show tendencies to form clusters rather than undergo vowel interpolation, though even in these contexts we find greater entropy for low vowels than non-low vowels.

As noted above, our treatment of non-low and low vowels could technically be reversed. We could assume only non-low vowels to be present in underlying representations (e.g. /jipuɻt̪/ → [jipuɻat̪a]/), then measure the predictability of word-internal low-vowel interpolation. As noted by a reviewer, these calculations would in fact show low vowels to be substantially predictable, almost as much so as non-low vowels in our model. What this ultimately shows is that for most C_C contexts in Anindilyakwa, it is highly predictable whether a vowel will be interpolated, or not. For example given the consonant sequence /t̪ŋn̪t̪ŋ/ ‘sharp’, it is highly predictable that word-internal vowels will be interpolated as in [t̪VŋVn̪t̪Vŋ]. The dimension of unpredictability is largely a binary choice as to whether each of these vowels is either low or non-low. This means

Table 5: Predictability of low vowels in various C_C types

Context type	CaC count	CC count	$H(ə \sim \emptyset)$
Equal or increasing sonority	2495	3628	0.97
l_N	95	55	0.95
ʎ_T	147	162	0.99
M_T	159	106	0.97
r_N	34	52	0.97
ɻ_k	42	62	0.97
Y_L	275	135	0.91
Y_N	224	177	0.99
Y_T	197	68	0.82
Dorsal_labial	69	415	0.59
Homorganic nasal_stop	60	852	0.35
L_m	34	97	0.83
L_T	152	673	0.69
N_P	38	77	0.92
r_ŋ	28	177	0.58
ɻ_m	18	56	0.80
ɻ_T	30	151	0.65
WEIGHTED AVERAGE			0.86

that a redundancy-free encoding only needs to represent one vowel type or the other, since whichever type is not represented can be predictably interpolated according to C_C context. For example by specifying the low vowels, we can predictably insert non-low vowels in the remaining contexts that require a vowel, as in /t̪ŋan̪t̪ŋ/ ⟶ [t̪əŋan̪t̪əŋ]. The analysis pursued here selects low vowels for underlying representation, rather than non-low vowels, in keeping with the earlier proposal by Heath (2020). But besides evaluating Heath's analysis, there are two reasons to model word-internal epenthesis of the non-low vowels, rather low vowels. Firstly, at word-internal morphological boundaries it is non-low vowels that are epenthesised, rather than low vowels. Secondly, it is non-low vowels that are cross-linguistically attested as word-internal epenthetic segments, rather than low vowels.

We have thus far shown that the presence/absence of non-low vowels in the Anindilyakwa lexicon is to a great extent predictable, based on wordlist data. Not only are there certain C_C contexts in which non-low vowels are predictably interpolated (as observed in earlier works), but there are also C_C contexts in which non-low vowels are predictably absent. We used an IT measurement to quantify the information value of non-low vowels, and compare this to the much greater information value of low vowels in the resulting lexical representations. In the following sections we will investigate how these findings on vowel predictability align with patterns of vowel omission in speech production data.

5 Variable omission of predictable vowels in production data

In this section we turn to speech production data, focusing in particular on the extent to which vowels are omitted. As noted above, previous research suggests that highly predictable segments should be articulated with reduced phonetic cues, including complete omission, in comparison to more informative segments (Lindblom 1990; Hall 2009, Hall et al. 2018: 119). We here present evidence that this is indeed the case with respect to Anindilyakwa vowels.

Since predictable segments are expected to undergo phonetic reduction, up to and including omission, one might also investigate whether predictable vowels in Anindilyakwa exhibit more gradient reduction in comparison to unpredictable vowels. However, this approach is confounded by other phonetic factors. Epenthetic final *-a* vowels in Anindilyakwa, although they are completely predictable, tend to be phonetically long when they are present. However, this observation is difficult to disentangle from the fact that final lengthening is a widespread, if not universal, property of speech (Fletcher 2010; Seifart et al. 2021). As for the predictable non-low vowels, these *do* exhibit shorter durations in comparison to the unpredictable word-internal low vowels. In preliminary analyses, the average durations of word-internal low vowels are around 100ms, while non-low vowels are around 50–60ms (Billington et al. 2022). However in this case, the shorter duration of non-low vowels cannot easily be disentangled from the fact that low vowels are cross-linguistically longer than non-low vowels (Lindblom 1967). In addition, there are likely to be influences of word-level prominence patterns on vowel realisation, but these are difficult to take into account based on current knowledge of Anindilyakwa prosody. To avoid these confounds, the analysis below focuses on the complete omission of vowels, rather than gradient phonetic reduction.

5.1 Data and method

All data for this study were collected using field elicitation of Anindilyakwa utterances, comprising a combination of picture prompts and spoken English prompts. Sets of around 50–100 utterances were collected in this way from seven Anindilyakwa speakers, and audio recorded. The same prompts were used for each speaker, so that their data is largely comparable and contains repeats of many of the same lexical items; however speakers were not coerced into using specific sentence frames or lexical items, so there is also substantial diversity in how they chose to translate the prompts.

The speakers' ages range from approximately 25 to 80. Five are women and two are men. They all speak Anindilyakwa as their main daily language, and all are multilingual in Kriol, English and other regional languages (especially Wubuy and Yolngu Matha). Sentences were transcribed by the first author, then converted into an EMU-SDMS hierarchical database (Winkelmann et al. 2017) by the third author, following automated phone segmentation via WebMAUS (Kisler et al. 2017) using the Australian Aboriginal Language model and by manual correction by all three authors. The resulting database contains a total of 493 sentences and 5668 vowel tokens. Non-low vowel tokens (N=1918) were labelled [i, ə, u] according to perceived vowel quality, but as in the preceding section they are treated as a single class for the purposes of the present analysis focusing on whether they are produced or not.

5.2 Omission of final *-a*

Beginning with the category of low vowels [a, ɛ] (N=3674), there is evidence for an association between predictability and omission. As we saw above, the presence of word-internal low vowels is lexically informative, and must therefore be maintained in representations such as /jpɻaṯ/ 'wallaby'. We should therefore expect these vowels to retain robust cues, and be reliably present, and this is supported by the production data. Our data includes 2071 instances where word-internal low vowels are expected to occur, and we did not identify any instances of vowel omission in these contexts.

The opposite situation should occur with word-final *-a*, which as we saw is completely predictable in the orthographic wordlist data. Indeed, in the speech production data, we find that final *-a* is quite often omitted (supporting informal observations by Leeding 1989: 139). It is consistently present in sentence-final words (N=493),[8] but it is only present in 67% of non-final words (total N=1110).

[8]We have observed just a handful of exceptions, for example (RaLa_nest_MED_01; RaLa_river_-FIN_01).

Leeding associates the omission of final *-a* with an initial low vowel in the following word. While we do find many such examples in our data, as for *kuʈanəm* followed by [ɛ] in (3), we also find many examples of *-a* omission where the following word is consonant-initial, as for *niŋɛn* followed by [k] in the same example. (Example (3) also exhibits word-internal non-low vowel omissions; to be discussed below).

(3) niŋɛn kuʈanəm ɛnmaɲc ɛjka
niŋɛn k-wʈa-n-m ɛn-maɲc ɛjk
1S IRR-climb-NPST-MUT PROX.NEUT-LOC tree
'I'm going to climb this tree.' (JoMa_tree_FIN_01)

In our production data *-a* omission is especially frequent at the boundary between a demonstrative and the following noun (as observed by Heath 2020), for example *t̪ak* in (4). This suggests that its presence may be disfavoured within some constituent types such as NPs, which are likely to form a phonological phrase.

(4) [t̪ak t̪urkʷarəkʷa]$_{NP}$ təŋan̪t̪əŋa
t̪ak trkʷarkʷ t̪ŋan̪t̪ŋ
DIST.FEM spear.grass sharp
'That spear-grass is sharp.' (RaLa_speargrass2_MED_01)

The fact that final *-a* is most consistently present utterance-finally, and is frequently absent within NPs, which appear to form phonological phrases, suggests that it may have a prosodic boundary-marking function. Its presence likely also interacts with pause phenomena. Further research would be required to disentangle how final *-a* interacts with prosodic or syntactic phrasing, since at present there is no description of syntax-prosody mapping in Anindilyakwa. However, the crucial point for our study is that final *-a* is freely omissable because it does not carry any lexical information – its main functions appear to relate to boundary marking, and it makes no contribution to distinguishing lexemes.

5.3 Omission of non-low vowels

Finally, we turn to non-low vowels [i, ə, u] in our speech production data, which as we saw above are largely predictable in the wordlist data, and therefore can be expected to be more subject to omission compared to word-internal low vowels. This is indeed the case in the production data. We focus on C_C types that are identified from the wordlist as either predictably interpolated, or as lexically unpredictable (Section 4.1). Our data contains 2582 tokens of these C_C types, and in 28% of these the non-low vowel is omitted.[9] We find that non-low vowels are often omitted both morpheme-internally (where the wordlist suggests that they are largely predictable), and at morpheme boundaries (where all sources concur that they are totally predictable).

As we saw above, C_Cs with equal or increasing sonority are attested with almost exceptionless non-low vowel interpolation in the wordlist data. But in the production data we find that vowels are variably omitted in these contexts, including in lexemes that the wordlist attests with vowel interpolation. Examples of vowel omission in increasing-sonority contexts can be seen for [pm] in (5) and [kʷm] in (6).

(5) jikarpma
jkarp-m
woomera-INSTR
'with a woomera' (EdMa, woomera_FIN_01)

(6) jɛʎukʷmaɲca
jɛʎkʷ-maɲc
rain-LOC
'in the rain' (JuLa, rain)

We also find examples of increasing-sonority clusters that involve a labial stop or nasal followed by either ɻ or l̪, as in [pɻ] in (7), [pl̪] in (8), and [ml̪] in (9).

(7) jipɻat̪a
jpɻat̪
'wallaby' (JoMa_kangaroo_MED_01)

(8) mɛmɛrpa məl̪arkpl̪al̪a
mɛmɛrp m-l̪arkpl̪al̪
calf fem-thin
'thin calves' (EdMa_lowerleg2_MED)

[9] Out of these 723 instances of non-low vowel omission, approximately 50 instances show marginal evidence of vocalic-like acoustic material (< 10ms in duration) intervening between consonant segments. We annotated these as vowel omission, since they are auditorily difficult to distinguish from transitional phenomena in clusters. This does not materially affect our claims, since the marginal tokens account for less than 10% of non-low vowel omissions, and in any case are quite consistent with our hypothesis regarding predictability and phonetic reduction.

(9) at̪aʎəml̪aŋa
at̪aʎm-l̪aŋʷ
river-GEN
'across the river' (KaMa, river)

There are also many examples of nasal-nasal, nasal-lateral or lateral-lateral clusters, as in (10)–(13). In (13) we also observe that the palatal glide attested in the wordlist form *micijɛʎa* 'beach', is omissible in speech.

(10) jinmamuwa
jnmamw
'egg' (CoMa_egg_FIN_01)

(11) niŋɛn ŋmarəŋnam
niŋɛn ŋ-ma-rŋk-na-m
1S 1S-FEM-see-NPST-MUT
'I see it (FEM).' (JoMa_road_FIN_01)

(12) ɛnl̪aŋʷa
ɛn-l̪aŋʷ
PROX.NEUT-GEN
'for this (NEUT)' (CoMa_shoulder_MED_01)

(13) micɛʎʎaŋʷuja
mcjɛʎ-l̪aŋʷ-wj
beach-GEN-COM
'along the beach' (CoMa_beach_FIN_01)

We also find widespread non-low vowel omission in C_Cs where the second consonant is a glide, as in [rj] in (14) and [jw] in (15). In some of these cases it can be difficult to define whether a vowel is phonetically present or not, given the acoustic similarities between vowels and glides.[10]

(14) mijɛrja
mjɛrj
'nest' (KaMa_nest_FIN_01)

(15) ŋajwa
ŋajw
'I…' (JuLa_water_FIN_01)

[10] In tokens of this type, we annotate vowels based on either intensity or formant changes, either of which suggests a vocalic interval that can be distinguished from the glide.

Overall, these data present an interesting situation where Anindilyakwa can be said to have very strict cluster constraints, based on citation forms in the wordlist. But in natural speech samples, even in a relatively careful register used in elicitation, Anindilyakwa is much more permissive of clusters.

We turn now to decreasing-sonority C_Cs, which in the written wordlist data comprise a complex mixture of types that are predictably interpolated, and unpredictable types that have lexemes listed both with and without vowels. Our speech production data is not extensive enough to investigate these subtypes in detail, but overall we find that decreasing-sonority C_Cs often appear as clusters. For example, we find clustering in non-homorganic nasal-stop sequences such as [mkʷ] in (16) and [mt̪] in (17).

(16) nuŋumkʷul̪ama
nŋ-mkʷl̪a-m
1s-stay.NPST-MUT
'I'm staying' (JuLa_shelter2_FIN_01)

(17) mamt̪akpa
mamt̪akp
'tail' (KaBa_tail_FIN_01)

We also find clustering of liquids followed by stops (18) or nasals (19), though these again present some examples where it is difficult to define what counts as an intervening vowel (cf. van Egmond 2012: 25ff.):

(18) jɛʎkʷa
jɛʎkʷ
'rain' (JoMa_rain_MED_01)

(19) al̪ŋacira
a-l̪ŋacr
NEUT-tall
'tall (NEUT)' (CoMa_tree_MED_01)

In our analysis of the wordlist we described how specific decreasing-sonority C_C types have lexically unpredictable non-low vowels. These constituted 18% of all C_C tokens in the lexicon and contributed most of the 0.23 bits of weighted average entropy. But the production data shows that there is variable omission of non-low vowels in these contexts, including evidence of variable production of the same lexeme. We have also noted that in some decreasing-sonority C_C types involving liquids and approximants (e.g. r_ŋ, l_m, ɻ_m, r_j), it is difficult to define whether a brief intervening vowel is present or not. We note that these types are also strongly represented in the purportedly unpredictable C_C contexts in the lexicon (see Table 3). This suggests that the apparent lexical unpredictability of non-low vowels in these contexts could instead reflect phonetic variation. For example if a vowel is variably present or absent in lexemes such as /mɛrŋʷ/ → [mɛr(u)ŋʷa] 'yellow clay', and /aml̪rŋʷ/ → [amul̪ur(u)ŋʷa] 'heel', this might result in apparent lexical distinctions in the wordlist, which attests r_ŋʷ vowel interpolation in *mɛruŋʷa*, but clustering in *amul̪urŋʷa*. Thus there is reason to suspect that even the limited informativity of non-low vowels represented in the wordlist may be an over-estimate, though more extensive phonetic research is required to test this.

6 Discussion and conclusions

In this chapter we have investigated vowel predictability in Anindilyakwa, analysing types of consonant contexts in which non-low vowels are predictably present, predictably absent, or lexically specified as either present or absent. We used an IT approach to conceptualise and quantify the information content of these vowels, and to compare this to the word-internal low vowels, which are much more informative.

We then examined evidence for the proposed association between low information content and reduced phonetic realisation, focusing on complete omission of vowels. Our speech production data provides support for this proposal, since we observed that the low-information vowel types, namely non-low vowels, and word-final *-a* vowels, are quite frequently omitted even in a relatively careful register of elicited speech. This is in contrast to the highly informative, word-internal low vowels, which are never omitted in our production data.

Our findings support a theory of predictable vowels in terms of informational vacuity and phonetic facilitation. On the one hand, predictable vowels do not in themselves contribute to the phonological contrasts that distinguish an intended word from all the others in the lexicon. On the other hand, highly predictable vowels may nonetheless play a role in speech production and perception, as has been argued for epenthetic vowels in previous work (Côté 2000, Tily & Kuperman 2012). The non-low vowels of Anindilyakwa may be largely uninformative in their own right, but they nonetheless may play a role in the accurate transmission of their information-rich neighbouring consonants, as has been proposed for vowels in Australian languages more generally (Butcher 2006). On the level of phonological categories and contrasts, non-low vowels in Anindilyakwa are uninformative because they rarely contribute to distinguishing one lexeme from another. But on the level of phonetic articulation, they facilitate the transmission of cues to consonant place of articulation. In IT terms, this is an issue of channel capacity, where a certain amount of redundancy in signal encoding (such as the interpolation of uninformative vowels) can help to maintain transmission fidelity. It has been noted in previous work that Anindilyakwa words seem unusually long in general (Leeding 1989: 68), and this is exactly what we might expect in an encoding system that includes very low-information segments (Nettle 1995, 1998).

As for the word-final epenthesis of *-a*, previous analyses already showed that this is completely non-contrastive because it can occur on all words, and therefore is completely uninformative with respect to the lexicon. But final *-a* likely has other functions. For one thing, it provides a sonorous substrate for intonational boundary tones, expressing pragmatic intent (a different kind of 'information'). In some contexts, lengthening of clause-final vowels is suggested to carry aspectual information (Bednall 2019). In our observations of low-vowel omission we noted that it is almost never omitted utterance-finally, and is omitted most frequently within noun phrases. This suggests the possibility that final *-a* may also help signal phrasal structure, though further research would be required to disentangle syntactic and prosodic phrases.

6.1 Predictability in gradient lexical representations

As pointed out near the beginning of this chapter, one advantage of IT is that it can capture gradient as well as categorical effects. Yet throughout this chapter we have associated the predictability of non-low vowels with their categorical omission from redundancy-free lexical representations as in /jpɻat̪/ → [jipuɻat̪a]. Other gradient approaches to phonology, such as exemplar theory (Bybee 2000, Pierrehumbert 2001), instead propose phonetically detailed, gradient lexical representations. In these approaches, each word has its own phonetically detailed representation, rather than being merely a conjunction of abstract phonological segments or features. This explains why phonetic variation shows lexically specific effects (e.g. Jurafsky et al. 2002, Walker 2012). Furthermore, the memory traces of specific phonetic tokens are said to be stored as 'exemplars', so that a lexical representation is a collective of all the exemplars experienced for that word. In this approach to phonology, phonemes and features are not the basic building blocks, but instead are epiphenomena that reflect the convergence of lexical representations on similar gestures and timing patterns (Bybee 2000: 72).

While the focus of this study has been discrete phonological representations, we believe that both our conceptual framework and our findings are also compatible with gradient lexical representations. For example, if we assume phonetically gradient representations along the lines of the 'gestural scores' used in articulatory phonology (Browman & Goldstein 1992), the IT approach would imply that certain gestures are more informative than others. In a word like [mamut̪akpa] 'tail', the vowel [u] would be present in the lexical representation as a (weak) vocalic gesture; but crucially, this gesture would be represented as having low surprisal with respect to the rest of the lexicon. And if lexical representations encompass probabilistic variation over phonetically distinct exemplars, then this would be reflected in a distribution of exemplar tokens encompassing both presence and absence of the weak vocalic gesture, i.e. [mamut̪akpa ~ mamt̪akpa], as observed in our production data. An IT-enriched exemplar theory would make the connection between this phonetic variation and the low informativity of the vocalic gesture (e.g. van Son & van Santen 2005, Hall 2009, Cohen Priva 2015, 2017, Shaw & Kawahara 2017, Hall et al. 2018). It is beyond the scope of this discussion to flesh out how exactly gestural scores and exemplar distributions might be enriched by representations of gestural surprisal, but the approaches do appear to be fundamentally compatible.

In this chapter we have discussed redundancy-free representations such as /mamt̪akp/, in which non-low vowels are treated as epenthetic. However, if we instead assume that lexical representations are phonetically gradient exemplar distributions, then we do not need to make an analytic choice between deletion and epenthesis. Instead, both vowel presence and absence are possibilities within a continuous space of gestural timing and magnitude. Furthermore, this approach supports the analysis of vowel reduction as a dynamic process, and one that may lead to changes over time (Wedel et al. 2013, Hall et al. 2018).

6.2 Directions for further research

The current study leaves open many questions. One of the most important is the matter of vowel quality in Anindilyakwa non-low vowels. In this study we have focused solely on how consonants condition vowel presence/absence, while setting aside the issue of how the quality of non-low vowel phones is conditioned by consonantal place-of-articulation. Investigation of this issue may further demonstrate the contextual predictability of these vowels, or it may reveal dimensions of lexical specificity that were not accessible in the current study. Detailed acoustic analyses would also further inform our understandings of gradient patterns in Anindilyakwa vowel production, as would analyses that are able to take into account the frequency of phones and lexemes based on natural speech data. Another issue that requires further research is the role of prosodic prominence, both in terms of clarifying prosodic patterns in Anindilyakwa, and examining how these interact with segmental phenomena.

Finally, the interpretation of our results will be greatly enhanced by comparative research on other languages. We have measured degrees of predictability in Anindilyakwa non-low vowels, comparing this to the low vowels. This has been partly motivated by debates about phonemic status in the Anindilyakwa vowel system, and the language's unusual structural patterns relative to the areal typology. But would equivalent measurements for other languages show that Anindilyakwa is in fact unusual in terms of vowel predictability? Do non-low vowels typically exhibit more contextual predictability than their louder and longer counterparts in the low vowel space? A good place to begin such research would be to compare data from other Australian languages, many of which have roughly similar segmental inventories to Anindilyakwa, but in which there has been no comparable debate about which vowels are lexical or epenthetic.

Acknowledgments

We wish to thank all the Anindilyakwa speakers who taught us about their language, and especially the seven who spoke sentences for this study: Ramesh Lalara, Kathleen Mamarika, Katelynn Bara, Edith Mamarika, Judy Lalara, Coleen Mamarika and Joel Marawili. We also thank staff at the Groote Eylandt Language Centre – James Bednall, Brighde Collins and Carolyn Fletcher – who provided generous support and hospitality to the first author in his field visits. Kathleen Currie Hall and two anonymous reviewers gave us valuable comments on an earlier draft. We gratefully acknowledge funding support from the University of Melbourne Research Unit for Indigenous Language, and the Australian Research Council grant DE180100872.

References

Bednall, James. 2019. *Temporal, aspectual and modal expression in Anindilyakwa, the language of the Groote Eylandt archipelago, Australia.* Australian National University & Université Paris Diderot. (Doctoral dissertation).

Bednall, James. 2021. Identifying salient aktionsart properties in Anindilyakwa. *Languages* 6(4). 164.

Billington, Rosey, John Mansfield & Hywel Stoakes. 2022. Acoustic and durational characteristics of Anindilyakwa vowels. In Rosey Billington (ed.), *Proceedings of the 18th Australasian International Conference on Speech Science and Technology*, 66–70. Canberra: Australasian Speech Science & Technology Association.

Breen, Gavan & Veronica Dobson. 2005. Central Arrernte. *Journal of the International Phonetic Association* 35(2). 249–254.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3-4). 155–180.

Butcher, Andrew. 2006. Australian Aboriginal languages: Consonant-salient phonologies and the 'place of articulation imperative'. In Jonathan M. Harrington & Marija Tabain (eds.), *Speech production: Models, phonetic processes and techniques*, 187–210. New York: Psychology Press.

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 65–85. Stanford: Center for the Study of Language & Information.

Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology* 6(2). 243–278.

Cohen Priva, Uriel. 2017. Informativity and the actuation of lenition. *Language* 93(3). 569–597.

Côté, Marie-Hélène. 2000. *Consonant cluster phonotactics: A perceptual approach*. Massachusetts Institute of Technology. (Doctoral dissertation).

Cover, Thomas A. & Joy A. Thomas. 2002. *Elements of information theory*. 2nd edn. London: Wiley.

Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The handbook of phonetic sciences*, 2nd edn., 521–602. Hoboken: Wiley.

Goldsmith, John A. 2000. On information theory, entropy, and phonology in the 20th century. *Folia Linguistica* 34(1). 85–100.

Hall, Kathleen Currie. 2009. *A probabilistic model of phonological relationships from contrast to allophony*. Ohio State University. (Doctoral dissertation).

Hall, Kathleen Currie. 2013a. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215–275.

Hall, Kathleen Currie, Elizabeth Hume, Florian T. Jaeger & Andrew Wedel. 2018. The role of predictability in shaping phonological patterns. *Linguistics Vanguard* 4(s2). 20170027.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2013b. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics* 41(2). 133–143.

Hartley, Ralph V. L. 1928. Transmission of information. *Bell System Technical Journal* 7. 535–568.

Heath, Jeffrey. 2020. Draft grammatical sketch of Anindilyagwa (Groote Eylandt, Australia). ms, Canberra.

Hockett, Charles F. 1967. The quantification of functional load. *Word* 23(1-3). 300–320.

Hume, Elizabeth, Kathleen Currie Hall, Andrew Wedel, Adam Ussishkin, Martine Adda-Dekker & Cédric Gendrot. 2011. Anti-markedness patterns in French epenthesis: An information-theoretic approach. *Annual Meeting of the Berkeley Linguistics Society* 37(1). 104–123.

Jurafsky, Daniel, Alan Bell & Cynthia Girand. 2002. The role of the lemma in form variation. *Laboratory Phonology* 7. 3–34.

Kisler, Thomas, Uwe Reichel & Florian Schiel. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347.

Leeding, Velma. 1989. *Anindilyakwa phonology and morphology*. University of Sydney. (Doctoral dissertation).

Lindblom, Björn. 1967. Vowel duration and a model of lip mandible coordination. *Speech Transmission Laboratory Quarterly Progress and Status Report* 8(4). 1–29.

Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling (NATO ASI series)*, 403–439. Dordrecht: Springer.

Nettle, Daniel. 1995. Segmental inventory size, word length, and communicative efficiency. *Linguistics* 33(2). 359–367.

Nettle, Daniel. 1998. Coevolution of phonology and the lexicon in twelve languages of West Africa. *Journal of Quantitative Linguistics* 5(3). 240–245.

Parker, Jeff. 2015. Solving Russian velars: Palatalization, the lexicon and gradient contrast utilization. *Slavic and East European Journal* 59(1). 70–90.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137–157. Amsterdam: Benjamins.

Seifart, Frank, Jan Strunk, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nikolaus P. Himmelmann & Balthasar Bickel. 2021. The extent and degree of utterance-final word lengthening in spontaneous speech from 10 languages. *Linguistics Vanguard* 7(1). DOI: 10.1515/lingvan-2019-0063.

Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27(3). 379–423.

Shannon, Claude E. 1951. Prediction and entropy of printed English. *Bell System Technical Journal* 30. 50–64.

Shaw, Jason A. & Shigeto Kawahara. 2017. Effects of surprisal and entropy on vowel duration in Japanese. *Language and Speech.* 62(1). 80–114.

Stokes, Judith. 1981. Anindilyakwa phonology from phoneme to syllable. In Bruce Waters (ed.), *Work papers of SIL-AAB series A , Australian phonologies: Collected papers*, vol. 5, 139–181. Darwin: SIL.

Tily, Harry & Victor Kuperman. 2012. Rational phonological lengthening in spoken Dutch. *The Journal of the Acoustical Society of America* 132(6). 3935–3940.

van Egmond, Marie-Elaine. 2012. *Enindhilyakwa phonology, morphosyntax and genetic position.* University of Sydney. (Doctoral dissertation).

van Egmond, Marie-Elaine & Brett Baker. 2020. The genetic position of Anindilyakwa. *Australian Journal of Linguistics* 40(4). 492–527.

van Son, R. J. J. H. & Jan P. H. van Santen. 2005. Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication* 47(1–2). 100–123.

Waddy, Julie Anne. 1989. Draft Anindilyakwa dictionary.

Walker, James A. 2012. Form, function, and frequency in phonological variation. *Language Variation and Change* 24(3). 397–415.

Wedel, Andrew, Scott Jackson & Abby Kaplan. 2013. Functional load and the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech* 56(3). 395–417.

Winkelmann, Raphael, Jonathan Harrington & Klaus Jansch. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 45. 392–410. DOI: 10.1016/j.csl.2017.01.002.

Chapter 4

The patterning of epenthesis in Urban Hijazi Arabic

Hassan Bokhari

This paper accounts for the different types and motivations of epenthesis in Urban Hijazi Arabic, such as syllable structure-driven epenthesis (SylE) and sonority-driven epenthesis (SonE). It also analyzes the default quality of the epenthetic vowel in SonE and SylE related to the type of the prosodic unit, syllable, or foot, specifically whether the epenthetic vowel is in the head syllable of the foot or the weak syllable. Beyond that, it analyzes the correlation between the quality of the nondefault SonE epenthetic vowel in final rising sonority coda clusters and the place features of the stem vowel or coda consonants.

1 Introduction

Urban Hijazi Arabic (UHA), also known as Makkan Arabic, is a prominent dialect spoken in the western region of Saudi Arabia, including the cities of Makkah, Jeddah, Madinah, to a lesser extent, Taif, and in some cities along the northwestern coast of Saudi Arabia.

UHA recognizes two major patterns of vowel insertion, both of which are incorporated into the syllabic structure of the word, making them epenthetic, as opposed to intrusive (Hall 2024 [this volume], Bellik 2024 [this volume]). These two types of epenthesis appear to be unique to this particular dialect since most other Arabic dialects typically insert the default vowel [i] (Farwaneh 2017: 102). These two patterns are syllable structure-driven epenthesis (SylE) where the default epenthetic vowel is [a], and sonority-driven epenthesis (SonE) where the default epenthetic vowel is [i]. In both SylE and default SonE there is a relationship between the type of prosodic unit where epenthesis occurs and the quality

Hassan Bokhari. 2024. The patterning of epenthesis in Urban Hijazi Arabic. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 85–104. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264534

of the default epenthetic vowel which will be discussed in more depth in Sections 2 and 3. Furthermore, the study analyzes the relationship between the stem vowel or one of the last two consonants of a CVCC word-final syllable, which dictates the quality of the nondefault SonE vowel in word-final rising sonority coda clusters in terms of the features CORONAL, DORSAL, and PHARYNGEAL.

Within the typology of syllable structure of Arabic dialects (Kiparsky 2003, Watson 2007, Broselow 2018), UHA is considered one of the CV dialects (or 'onset dialects' in Broselow's terminology) in which epenthesis occurs to the right of the unsyllabified consonant. For example, underlying /katab-t-l-u/ becomes [ka.tab.ta.lu] 'I wrote to him', with an epenthetic vowel to the right of /t/, forming the onset of the vowel.[1]

While Abu-Mansour (1987) and Kabrah (2004) have also studied epenthesis in the dialect, this study provides a novel, detailed analysis that differentiates two major types of epenthesis in UHA. The study also accounts for several systematic cases of consonant-vowel harmony in which the epenthetic vowel is not the default vowel but harmonizes with the [PLACE] feature of the consonant of a potential CVCC. The source of data analyzed in this paper is from Bokhari (2020).

This paper is organized as follows. Section 2 analyzes SylE and default SonE. Section 3 analyzes several types of non-default SonE, and Section 4 discusses some of the findings and concludes the paper.

2 SylE and default SonE

De Lacy (2006) provides several universal sonority-based constraint hierarchies and relates them to different positions of prosodic constituents. These sonority-based constraint hierarchies provide the right tool to set up the analyses of both types of default epenthesis. He argues that variation in the quality of the epenthetic vowel across languages can be analyzed as the result of competing constraints, imposed by different Designated Terminal Elements (DTEs). DTEs refer to the head of a given prosodic unit, such as a mora, syllable, or foot. He differentiates between the head and non-head positions of these constituents, which are the DTEs (also abbreviated as Δ) and non-Designated Terminal Elements (non-DTEs, abbreviated as $-\Delta$), respectively. According to De Lacy, universally, low vowels, which are higher in sonority than other vowels, tend to be favored as

[1] The final consonant of the CVCC syllable is extrametrical and does not bear any moraic weight. For example, underlying /katab-t/ surfaces as [ka$^{\mu}$.ta$^{\mu}$b$^{\mu}$t] with an extrametrical [t]. Therefore, UHA does not allow mora sharing.

epenthetic vowels in DTE positions, i.e., the head position of the prosodic constituent, whereas high peripheral vowels [i, u], tend to be epenthesized in non-DTE positions – that is, in unstressed syllables, moras, or feet.

According to De Lacy (2006: 305), vowel epenthesis in Arabic dialects cannot be accounted for in terms of DTE constraints alone. In UHA, however, the DTE of the prosodic constituents appears to be in some way connected with the sonority of the epenthetic segments, in which the low sonority vowel [i] is epenthesized in the non-head position of the foot in sonority-driven epenthesis, whereas the high-sonority [a] is epenthesized in the DTE position of the syllable in syllable-driven epenthesis. Since UHA has two major types of epenthesis, SylE epenthesis and default SonE epenthesis, it is worth elaborating and differentiating between the two types of epenthesis. I start by elaborating on SylE epenthesis then explain SonE epenthesis and summarize some issues related to it.

2.1 Syllable Structure-Driven Epenthesis (SylE)

Regarding the syllable-driven epenthetic vowel [a], this vowel, which is higher in sonority than any other vowel in UHA, is inserted as a way of strengthening a weak degenerate syllable, which consists of only a single consonant as a syllable onset. This also can be accounted for with the constraints proposed by De Lacy (2006: 68), in which this vowel ([a]) represents the DTE of the syllable. This epenthetic vowel could fall in DTE or non-DTE position of the foot.[2] In other words, the DTE of the syllable takes priority, in which the high sonority vowel fills the nucleus of the degenerate syllable. The trigger of this epenthetic vowel is the syllabic structure, in which an onset with an empty nucleus is not allowed in the dialect. Consider /ka.tabt.lu/ 'I wrote for him', which becomes [ka.tab.ta.lu], not *[ka.tab.t.lu], and /baːb.na/ 'our door', which becomes [baː.ba.na], not *[baː.b. na]. The vowel [a] is inserted after the [t] in the first example, which forces resyllabification of the [t] into the onset of the new syllable. The same is true in the second example, when [a] forms a new syllable with the preceding [b] in the process of resyllabification. Thus, the constraint ranking for this type of epenthesis is $*\Delta\sigma\leq\{i, u\} \gg *\Delta\sigma\leq a$, where $*\Delta\sigma\leq\{i, u\}$ is defined as the Designated Terminal Elements of a syllable may not be less sonorant than or equal to [i] or [u], and where $*\Delta\sigma\leq a$ is defined as the Designated Terminal Elements of a syllable may not be less sonorant than or equal to [a].

A potential OT analysis for the SylE epenthesis would rank $*\Delta\sigma\leq\{i, u\}$, Nucleus, which requires each syllable to have a vowel nucleus, and *ComplexOn-

[2] See the tableau in (9).

SET, which bans onset clusters, above *Δσ≤a. This analysis is beyond the scope of this paper and is left for future research.

De Lacy (2006) relates the sonority of the vowel to the head of the syllabic constituent; however, his proposal has a wider scope, in which he considers several universal sonority hierarchies of segments and how they interact with the head or non-head position of different constituents.

2.2 Sonority-driven default epenthesis (SonE)

In UHA, the sonority-driven default epenthetic vowel [i] is inserted to break up a potential word-final CC cluster of rising sonority when the quality of the vowel is not determined by high vowel spreading, a pharyngeal/laryngeal, coronal, or dorsal segment. Consider the data in Table 1, in which [i] is inserted to break up the rising-sonority cluster.

Table 1: Default Sonority-Driven Epenthesis

Underlying	Gloss	Surface	Possessive.3SG.MASC
a. /lakm/	‘punching’	[la.kim]	[lak.mu]
b. /tˤagm/	‘set (of things)’	[tˤa.gim]	[tˤag.mu]
c. /ʃamʕ/	‘wax’	[ʃa.miʕ]	[ʃam.ʕu]
d. /fagʕ/	‘type of mushroom’	[fa.giʕ]	[fag.ʕu]
e. /ʃatm/	‘cursing’	[ʃa.tim]	[ʃat.mu]
f. /nad͡ʒm/	‘star’	[na.d͡ʒim]	[nad͡ʒ.mu]
g. /waʃm/	‘tattoo’	[wa.ʃim]	[waʃ.mu]

As can be noted from the data, epenthetic [i] is not determined by the nature of the surrounding consonants. Regarding (1e–f), the [t] and [d͡ʒ] are not the trigger of [i]-epenthesis, even though they agree with the vowel [i] in the feature [coronal].[3] This is because the epenthetic vowel is required to harmonize in coronality with a following consonant if it is coronal as will be analyzed in section 3, and not with the preceding consonant. Therefore, [i] is epenthesized when there is no harmony requirement between it and the segments in the coda cluster. In this dialect, the only coda cluster permitted by the syllabic structure is the coda in a CVCC final syllable that does not have rising sonority. This monosyllabic

[3]The coronality of [i] is still a topic of discussion. See Clements (1991) and Hume (1992) for arguments that support the coronality of [i].

word form also consists of a trochaic foot, in which the stem vowel is the head of the foot, even if the coda cluster of this syllable violates the sonority requirement and receives an epenthetic vowel. This epenthetic vowel is never stressed, and it falls in the non-head position of the foot – that is, the unstressed part. Therefore, adopting De Lacy's DTE model mentioned above, I can determine the type of default epenthetic vowel in this position, following de Lacy's constraint ranking in the non-DTE position of the foot, by having *−ΔFt≥a outrank *−ΔFt≥{i,u}, where *−ΔFt≥a is defined as the head of the Non-designated Terminal Element may not be less than or equal in sonority to the low vowel [a], and where *−ΔFt≥{i,u} means that the head of the non-Designated Terminal Element may not be less than or equal in sonority to the high vowels [i] and [u]. This leaves us with the two high vowels [i, u] as potential candidates for a default vowel. On the basis of the Place of Articulation hierarchy, [u] is universally more marked than [i] (Lombardi 1995). Consider Example (1), which presents the universal Place of Articulation hierarchy.

(1) Universal Place of Articulation Hierarchy (De Lacy 2006)
    *[DORSAL] ≫ *[LABIAL] ≫ *[CORONAL] ≫ [PHARYNGEAL]

Therefore, I can say that *[Dors] outranks *[Cor]. By these two different rankings by De Lacy and Lombardi, I reach the conclusion that the most appropriate default epenthetic vowel is [i] in sonority-driven epenthesis in UHA. Before providing the OT analysis of default SonE, it is worth elaborating on the theoretical framework used in the OT analysis of the word-final coda clusters.

This study employs the Split-Margin Theory to analyze coda clusters in UHA (Baertsch 2002, Baertsch & Davis 2009). This approach to the syllable provides a framework for analyzing the behavior of coda clusters in terms of sonority. Under this theory, the onset and coda positions in the syllable (i.e., the syllable margins) are each optionally split into two positions, M_1 and M_2, where M_2 is the position closest to the nucleus of the syllable in each margin, and M_1 is the position farthest from the nucleus. The Sonority Sequencing Principle (SSP) states that sonority is highest at the nucleus and lowest at the edges of a syllable, making M_1 a low-sonority position and M_2 a high-sonority position. Because the SSP also states that cross-linguistically, syllables prefer to begin with low-sonority segments and end with high-sonority segments, a singleton onset is M_1, while a singleton coda is M_2.

The M_1 position gives preference to low-sonority segments. When a *M_1 constraint is aligned with the sonority hierarchy, constraints avoiding high sonority will be universally highly ranked, and constraints avoiding low sonority will be universally lowly ranked. In UHA, $*M_1$ and $*M_2$ OT constraint hierarchies are inherently ranked following the UHA sonority scale. The M_1 position gives preference to low-sonority segments, while the M_2 position gives preference to high-sonority segments.

(2) $*M_1$/Vowel ≫ $*M_1$/Glide ≫ $*M_1$/ʕ ≫ $*M_1$/Liquid ≫ $*M_1$/Nasal ≫ $*M_1$/VcdFri ≫ $*M_1$/Obs
$*M_2$/Obs ≫ $*M_2$/VcdFri ≫ $*M_2$/Nasal ≫ $*M_2$/Liquid ≫ $*M_2$/ʕ ≫ $*M_2$/Glide ≫ $*M_2$/ Vowel

Now, let us consider the tableau below in (2).

Table 2: Default SonE [i]-Epenthesis

/lakm/ ‘punching’	*−ΔFt≥a	*[Dors]	*−ΔFt≥i,u	*[Cor]	*[Phar]
☞ a. [la.kim]			*	*	
b. [la.kam]	*!		*		*
c. [la.kum]		*!	*		

In this tableau (2), candidate (a) wins because it respects both *−ΔFt≥a and *[Dors] by epenthesizing a coronal vowel, even though it violates low-ranked *−ΔFt≥{i,u} and *[Cor]. Candidate (b) loses because the epenthetic vowel [a] violates *−ΔFt≥a; in addition, it violates *−ΔFt≥{i,u} and *[Phar], because [a] as a pharyngeal vowel is greater in sonority than [i] and [u]. Candidate (c) loses because it epenthesizes a dorsal vowel, violating the high-ranked *[Dors]; in addition, it violates *−ΔFt≥{i,u}. This tableau shows that whenever there is no harmony requirement between vowels or consonants and vowels, default [i] is epenthesized to break up a potential rising-sonority coda cluster.

In contrast, the motivation for SylE is the tendency of the dialect to avoid word-internal superheavy syllables by epenthesis and resyllabification, in which the last stray consonant of the word-internal CVVC/CVCC is resyllabified to form the onset of the default [a] epenthetic vowel. The outcome of this epenthesis is the new syllable, which is formed by the unsyllabified last consonant of the superheavy syllable and the epenthetic vowel, and this syllable is preceded by a heavy syllable (e.g., /baːb- na/ ⟶ [(ˈbaː).(ba̠.na)] ‘our door;’ /katab-t-l-u/ ⟶ [ka.(ˈtab).(ta̠.lu)] ‘I wrote for him’). In both types of epenthesis, stress location is preserved, even after epenthesis. Thus, I can conclude from the discussion above that the SylE epenthetic [a] is higher in sonority than the default SonE epenthetic [i], because [a], as the highest sonority vowel, forms the nucleus of the syllable with the stray consonant, i.e. the DTE of the syllable that is not the weak part of a foot, whereas the lowest sonority high front vowel [i] is epenthesized in the default SonE in the non-DTE position of the foot. Note that word-final foot extrametricality prevents penultimate stress in these words.

3 Non-default Sonority-Driven Epenthesis (non-default SonE)

After differentiating between two major types of epenthesis in UHA, this section analyzes non-default SonE, which operates whenever the coda cluster would exhibit a rising sonority profile. The quality of both stem vowels and the consonant in the coda cluster play a role in determining the quality of the epenthetic vowel, which breaks up a potential rising coda cluster. An underlying high vowel spreads its features to the epenthetic vowel on the surface in the high vowel spreading operation. [i] is epenthesized agreeing with final coronal consonants, except pharyngealized [rˤ]. [a], which is a pharyngeal vowel, is epenthesized agreeing in the feature [pharyngeal] with a preceding pharyngeal/laryngeal consonant and a following pharyngeal rhotic. In words with no medial laryngeal/pharyngeal consonant, [u] is epenthesized agreeing with the final pharyngealized rhotic [rˤ] for old generation speakers of the dialect. Several OT constraints are used in the analysis of non-default SonE, that includes constraints related to high vowel spreading, and consonant-to-vowel harmony. When both Dep, which militates against epenthesis, and Contig, which militates against separating two adjacent underlying segments in the surface form, are dominated by any of the Split-Margin constraints, the non-default SonE constraint ranking operates.

3.1 High Vowel Spreading

In UHA, when the stem contains a high vowel [i] or [u], the epenthetic vowel, which breaks up the potential rising-sonority word-final coda cluster, is the result of the autosegmental spreading of the underlying high vowel in the stem. In simpler terms, if the stem has the high vowel [i], the epenthetic vowel is [i], and if the stem vowel is [u], the epenthetic vowel is [u], as shown in the data in Table 3.[4]

[4]It is worth mentioning that the vowel inventory of UHA includes the vowels /i/, /iː/, /u/, /uː/, /a/, /aː/, and the mid vowels [ee] and [oo]. The mid vowels [ee] and [oo] are not underlying in the dialect and are phonologically derived.

Table 3: High vowel spreading in potential rising coda clusters in UHA

Underlying	Gloss	Surface	Possessive.3SG.MASC
a. /gidr/	'pot'	[gi.dir]	[gid.rˤu]
b. /ʔism/	'name'	[ʔi.sim]	[ʔis.mu]
c. /fiʕl/	'verb, action'	[fi.ʕil]	[fiʕ.lu]
d. /ħukm/	'verdict, ruling'	[ħu.kum]	[ħuk.mu]
e. /χuʃm/	'nose'	[χu.ʃum]	[χuʃ.mu]
f. /duɦn/	'fat'	[du.ɦun]	[duɦ.nu]

As illustrated in (3), the sonority-driven epenthetic vowel has the same quality as the stem vowel, because the stem vowel is high; however, when the stem vowel is low, consonant-to-vowel harmony can take effect. Otherwise, the default epenthetic vowel is [i] as discussed in the preceding section. I analyze the high vowel spreading epenthesis – loosely following the logic of Walker (2001) in her analysis of Altaic rounding harmony – as a process of autosegmental spreading: the [front] feature spreads to the epenthetic vowel if the stem contains a [high] [front] vowel, and the [back] feature spreads to the epenthetic vowel if the stem contains a [high] [back] vowel. In order to motivate the spreading of the [front] or [back] feature to the epenthetic vowel, the constraints in (3) are necessary.

(3) a. **Spread-[front] (adapted from Walker 2001):**
'For any vowel in a word linked to a [front] autosegment, that same [front] autosegment must also be associated to all other vowels in the word. Assign a violation for any [front] autosegment that is not associated to all vowels in the word.'

b. **Spread-[back] (adapted from Walker 2001):**
'For any vowel in a word linked to a [back] autosegment, that same [back] autosegment must also be associated to all other vowels in the word. Assign a violation for any [back] autosegment that is not associated to all vowels in the word.'

c. **Ident-IO(Vowel):**
'Let α be a vowel in the input and β be a correspondent of α in the output; then α and β have identical featural specifications. Assign a violation for any discrepant featural specification between α and β.'

d. **Uniform-[front]/[back] (adapted from Walker 2001):**
'A [front] or [back] autosegment may not be multiply-linked to vowels that are distinctly specified for height.'

Table 4: High back vowel spreading

/ʃukr/ 'thank'	*O2L1]σ	IdentIO(Vowel)	Spread-[back], Spread-[front]	Uniform-[front]/[back]	Dep	Contig
a. [ʃukr]	*!					
☞ b. [ʃu.kur^{ʕ}]					*	*
c. [ʃu.kir]			*!		*	*
d. [ʃi.kir]		*!			*	*
e. [ʃu.kor]				*!	*	*

The faithful candidate (a) loses, because it violates the high-ranked Split Margin constraint by exhibiting a rising-sonority coda cluster. Candidate (b) wins, because the high back specification spreads to the epenthetic vowel; therefore, both vowels are identical. Candidates (c) and (d) both fatally violate Spread-[back] in addition to Dep and Contig. For (c), the epenthetic vowel is a high front vowel, which does not harmonize with the stem vowel. Candidate (d) loses, because it violates Ident-IO(Vowel), even though the vowel features are shared by both syllables. Candidate (e) loses, even though the [back] feature spreads to the epenthetic vowel, because the height specifications of the vowels are different; therefore, it violates Uniform-[back].

Having provided the analysis for stem high vowel spreading in the previous tableau, now I turn to explain consonant-to-vowel harmony in stems with an underlying low vowel.

3.2 Coronal consonant-to-vowel harmony

In UHA, in underlying CaCC words, in which the stem vowel is a low vowel and the last consonant is coronal, the vowel [i] is epenthesized if the last two consonants would form a rising-sonority coda cluster. Consider the examples in Table 5 below.

Table 5: Coronal consonant-to-vowel harmony

Underlying	Gloss	Surface	Possessive.3SG.MASC
a. /makr/	‘cunning’	[ma.kir]	[mak.rˤu]
b. /saɦl/	‘valley’	[sa.ɦil]	[saɦ.lu]
c. /laħn/	‘melody’	[la.ħin]	[laħ.nu]
d. /χabz/	‘baking’	[χa.biz]	[χab.zu]
e. /lafzˤ/	‘word’	[la.fizˤ]	[laf.zˤu]

As can be noted in (5), all words ending with a coronal segment require the preceding epenthetic vowel to be [i] in order to match the coronality (frontness) between the consonant and the epenthetic vowel. In addition, according to Padgett (2011), universally, there is a harmonic effect between coronal consonants and front vowels. Note that all intermediate consonants in these words are non-coronal consonants. Therefore, this confirms that the trigger of coronal harmony is the last consonant in the word, since coronal consonant-to-vowel harmony is regressive, i.e. from the last consonant of the word to the preceding epenthetic vowel. This will be clear if we compare the data above with the default SonE in words such as [lakim], in which the second and third consonants of the word are noncoronal consonants. Thus, there is a need for an additional constraint to regulate the relationship between the last coronal segment of the word and the epenthetic vowel in the environment of sonority-driven epenthesis in a rising-sonority coda cluster.

(4) **AGREE-FEAT-CORONAL (AGREE-F-COR):**
Segments immediately preceding and tautosyllabic with a coronal consonant must agree with it in the feature [coronal]. Assign a violation for any segment in the output which immediately precedes and is tautosyllabic with a coronal consonant and does not share the feature [coronal].

The tableau in (6) provides the analysis of CaCC words in which the last consonant is a coronal and requires the immediately preceding epenthetic vowel to be the coronal [i].

Table 6: Agree-F-Cor in coronal-final CaCC Words

/ʃakl/ 'shape, appearance'	*O2L1]σ	Agree-F-Cor	Dep	Contig
a. [ʃakl]	*!			
☞ b. [ʃa.kil]			*	*
c. [ʃa.kul]		*!	*	*
d. [ʃa.kal]		*!	*	*

Candidate (a) loses because it violates the Split-Margin constraint by exhibiting a rising-sonority coda cluster in the output form. Candidate (b) wins because [i] harmonizes with [l] by agreeing in the feature [coronal]. Candidate (c) loses because [u], which is a dorsal vowel, does not agree with the following coronal consonant. In the same way, candidate (d) loses because the epenthetic pharyngeal [a] does not agree with the [l].

Having provided the analysis in which a final coronal consonant dictates the quality of the preceding epenthetic vowel in the process of non-default SonE, now I turn to pharyngeal and laryngeal consonant-to-vowel harmony, in which these consonants dictate that [a] will be the surface form of the epenthetic vowel. This vowel can be followed by a pharyngealized [rˤ], which also contains a pharyngeal feature in its segmental representation.

3.3 Pharyngeal consonant-to-vowel harmony

Before explaining the reason why the words in (7) below receive the low vowel [a], it is worth explaining the status of /r/ in Arabic. In Arabic, /tˤ ðˤ dˤ sˤ/ are the main emphatic consonants, yet /r/ also has an emphatic allophone based on the dialect and some phonological conditions. Younes (1993: 121) argues that the status of the emphatic [rˤ] is not fully established. According to him, emphatic [rˤ] causes lowering in adjacent vowels.

Additionally, Herzallah's 1990 representation of emphatics in Palestinian Arabic, including underlyingly emphatic /rˤ/, shows that these consonants have a secondary place of articulation, which itself has two components: pharyngeal and dorsal. However, in contrast to Herzallah's proposal that emphatic /rˤ/ is underlying in Palestinian Arabic, data demonstrate that for the majority of UHA speakers /r/ is underlyingly only plain (coronal) and it is pharyngealized next to emphatic, pharyngeal, or laryngeal segments; with regard to underlying pharyngealized /rˤ/, it is associated with some speakers of the older generation, as will be analyzed in Section 3.4.

In UHA stems with a low vowel in which a potential coda cluster contains a pharyngeal or laryngeal consonant followed by pharyngealized [rˤ], the epenthetic vowel that is inserted to avoid the surfacing of such a sonority-rising cluster is the vowel [a], which also has a [pharyngeal][5] component. The words in (7) include medial pharyngeal or laryngeal consonants followed by pharyngealized [rˤ]. These words receive a low epenthetic vowel in the output form in order to avoid a potential rising-sonority coda cluster.

Table 7: Potential coda clusters with pharyngeal consonant-to-vowel harmony

Underlying	Gloss	Surface	Possessive.3SG.MASC
a. /ʃahr/	'month'	[ʃa.ɦarˤ]	[ʃaɦ.rˤu]
b. /naɦr/	'river'	[na.ɦarˤ]	[naɦ.rˤu]
c. /maɦr/	'dowry'	[ma.ɦarˤ]	[maɦ.rˤu]
d. /baħr/	'sea'	[ba.ħarˤ]	[baħ.rˤu]

In UHA, /r/ is pharyngealized next to emphatics, pharyngeals, laryngeals, and low and back vowels; otherwise, it is only coronal (plain), as shown in (8).

Table 8: Potential coda clusters with plain [r] and pharyngealized [rˤ]

Underlying	Gloss	Surface	*Nisba* Adjective	*Nisba* Gloss
a. /fikr/	'thought'	[fi.kir]	[fik.ri]	'intellectual'
b. /ʃukr/	'thanking'	[ʃu.kurˤ]		
c. /baħr/	'sea'	[ba.ħarʕ]	[baħ.ri]	'naval, nautical, marine'

In data set (8)a, [r] is only coronal (i.e., not emphatic) because it is not preceded by a low or back vowel nor by a laryngeal or pharyngeal consonant. Therefore, the epenthetic vowel to break this rising-sonority coda cluster is the default [i]. In (8)b, [rˤ] is pharyngealized because it is adjacent to the high back vowel [u], which itself is the result of high vowel spreading from the stem vowel to the epenthetic vowel. In (8)c, the underlying coronal (plain) /r/ becomes pharyngealized [rˤ], because it is preceded by a laryngeal or pharyngeal consonant. Therefore, the epenthetic vowel to break up such a rising-sonority coda cluster is the low vowel [a], which itself agrees with the surrounding consonants in the feature

[5]According to Herzallah (1990) the low vowel [a] has a pharyngeal feature.

[pharyngeal]. I can conclude from (8) that the trigger of the low vowel insertion in rising-sonority coda clusters ending with pharyngealized [rˤ] is the pharyngeal and laryngeal consonants. In derived forms, when the *nisba* (adjectival) suffix /-i/ or the first-person possessive suffix /-i/ is attached to the /r/-final stem, the /r/ resyllabifies to form an onset for the syllable containing /-i/; therefore, it surfaces faithfully as the coronal (plain) [r].

Before starting the OT analysis of pharyngeal and laryngeal consonant-to-vowel harmony, it is necessary to present the definition for the constraint needed in this analysis.

(5) The Constraint Necessary for Pharyngeal Consonant-to-Vowel Harmony AGREE-FEAT-PHARYNGEAL (Agree-F-Phar): Segments immediately following and tautosyllabic with pharyngeal and laryngeal segments must agree with them in the feature [pharyngeal]. Assign a violation for any segment in the output which immediately follows and is tautosyllabic with a pharyngeal or laryngeal segment and does not agree with it in the feature [pharyngeal].

The new constraint will be used for the analysis of pharyngeal and laryngeal harmony in coda clusters including [rˤ].

Table 9: Pharyngeal and laryngeal consonant-to-vowel harmony

/baħr/ ‘sea’	*O2L1]σ	Dep	Contig	Agree-F-Phar
a. [baħr]	*!	*		
b. [baħrˤ]	*!			
☞ c. [ba.ħarˤ]			*	*
d. [ba.ħar]		*!	*	*

The faithful candidate (a) loses because it fatally violates the Split-Margin constraint; in addition, it violates Agree-F-Phar by preserving the plain [r] next to the pharyngeal consonant. This is because the underlying rhotic /r/ is [rˤ] if it is immediately adjacent to laryngeal and pharyngeal segments. Candidate (b) also loses, because it violates the Split-Margin constraint, even though the pharyngealized [rˤ] agrees with the preceding consonant in the feature [pharyngeal]. Candidate (c) is the winner, because it satisfies the undominated Split-Margin constraint; in addition, it satisfies low-ranked Agree-F-Phar, even though it violates Dep and Contig. Candidate (d) is like the winner, except the [r] is not pharyngealized, which creates a violation of the Agree-F-Phar constraint.[6]

This section has provided the analysis of those situations wherein a potential coda cluster includes a pharyngeal or laryngeal consonant followed by /r/, and how such a cluster is broken up by the appropriate [a] vowel, which also has a pharyngeal component.

3.4 Dorsal consonant-to-vowel harmony with /rˤ/

According to Youssef (2019), complex /r/ (i.e. pharyngealized [rˤ]) could have V-place [dorsal] as well as C-place [coronal]. He also indicates that pharyngealized [rˤ] shows pharyngeal constriction alongside dorsal lowering. Youssef's description of the pharyngealized /rˤ/ matches the feature representation of pharyngealized /rˤ/ and other pharyngeal obstruents given by Herzallah (1990). Therefore, I adopt the feature representation given by Herzallah (1990) in which /rˤ/ has primary [coronal] place, which is dominated by a C node, and secondary [dorsal] and [pharyngeal] places, which are dominated by a V node. The dataset in (10) below represents the uncommon realization of the rhotic in UHA, as produced by the older generation of UHA speakers, in which the rhotic is underlyingly pharyngealized.

Consistent with the assumptions of OT, I propose that the use of pharyngealized [rˤ] and agreement between this consonant and the preceding epenthetic [u] happens simultaneously in the output form rather than positing an intermediate stage in which /r/ becomes [rˤ] followed by insertion of [u] to agree with this segment, as illustrated in (10).

Table 10: [rˤ] consonant-to-vowel harmony

Underlying	Gloss	Surface	Possessive.3p.sg.masc	Possessive.1p.sg
a. /d͡ʒadrʕ/	'wall'	[d͡ʒa.durʕ]	[d͡ʒad.rʕu] 'his wall'	[d͡ʒad.ri] 'my wall'
b. /sˤagrʕ/	'falcon'	[sˤa.gurʕ]	[sˤag.rʕu] 'his falcon'	[sˤag.ri] 'my falcon'
c. /badrʕ/	'full moon'	[ba.durʕ]	[bad.rʕu] 'his full moon'	[bad.ri] 'my full moon'

Compare the data above with the parallel data point for the speakers of the younger generation for which the underlying form of the rhotic is plain /r/ in (11).

[6]Recall in section (2.1) that the vowel [a] falls in the independent position of the foot as in the word [ka.(ˈtab).(tạ.lu)] 'I wrote for him'. However, in the word [(ˈba.ħạrˤ)], [a] is epenthesized in the non-head position of the foot for a different reason. Here, it is inserted to agree with the [Phar] feature of the surrounding consonants.

Table 11: Underlying rhotic for speakers of the younger generation of UHA

Underlying	Gloss	Surface	Possessive.3P.SG.MASC	Possessive.1P.SG
a. /badr/	‘full moon’	[ba.dir]	[bad.rˤu] ‘his full moon’	[bad.ri] ‘my full moon’
b. /sˤagr/	‘falcon’	[sˤa.gir]	[sˤag.rˤu] ‘his falcon’	[sˤag.ri] ‘my falcon’
c. /d͡ʒadr/	‘wall’	[d͡ʒa.dir]	[d͡ʒad.rˤu] ‘his wall’	[d͡ʒad.ri] ‘my wall’

The data in (10) show words ending with /rˤ/ for the older generation, which receive [u] as the epenthetic vowel. This epenthetic vowel, regardless of its quality, breaks up the rising-sonority cluster, as is expected. When adding the 1st person possessive suffix [-i], [r] is not pharyngealized, because it forms the onset of the syllable which contains the vowel [i]. In the same way, when adding the 3rd person masculine possessive suffix [-u], the allophonic [rˤ] surfaces, because it falls in the onset position of the syllable containing the vowel [u]. From these data, when a speaker from the older generation has an underlying /rˤ/, it necessitates the insertion of the vowel [u] in the process of sonority-driven epenthesis, as both segments agree in the feature [dorsal].[7]

This leads us to propose a constraint that requires agreement between pharyngealized [rˤ] and the segment next to it.

(6) **Constraints motivating Dorsal Consonant-to-Vowel Harmony**

a. AGREE-FEAT-DORSAL-[rˤ] (Agree-F-Dors-[rˤ]):
Segments adjacent to and tautosyllabic with [rˤ] must agree with it in the feature [dorsal]. Assign a violation for any segment in the output which is adjacent to and tautosyllabic with [rˤ] and does not agree with it in the feature [dorsal].

b. Max-V-Place:
The V-Place feature (secondary place of articulation) associated with a consonant in the input must have a correspondent in the output. Do not delete a V-Place feature (secondary place of articulation) from a consonant. Assign a violation mark for any secondary place of articulation feature which is deleted from a consonant in the output.

[7]I propose that the disagreement between the generations' underlying forms of the rhotic can be understood as lexicon optimization, in which some output forms of the older generation with allophonic de-emphaticized [r] (e.g., the first person possessive suffix) are reinterpreted as the underlying form of the rhotic by the younger generation (Holt 2015: 547–548). This reconstructed underlying form of the rhotic has an emphatic allophonic variant when it meets the condition of being next to a dorsal vowel. This proposal is compatible with the Synchronic Base Hypothesis (Hutton 1996). See Bokhari (2020: section 5.6) for more details.

The tableau in Table 12 provides the analysis in which pharyngealized [rˤ] necessitates the epenthesis of [u] to match its dorsal feature.

Table 12: Dorsal harmony with pharyngealized [rˤ]

/d͡ʒadrˤ/ 'wall'	*O2L1]σ	Max-V-Place	Dep	Contig	Agree-F-Dors-[rˤ]	Agree-F-Cor
a. [ʤadrˤ]	*!				*	
b. [ʤadr]	*!	*				
☞ c. [ʤa.durˤ]			*	*		*
d. [ʤa.dur]		*!	*	*		*
e. [ʤa.dir]		*!	*	*		
f. [ʤa.dirˤ]			*	*	*!	

The faithful candidate (a) loses because it violates the Split-Margin constraint. It also violates Agree-F-Dors-[rˤ]. This is because the underlying form for the rhotic is /rˤ/ for older speakers of UHA. In the same way, candidate (b) loses because it also violates the Split-Margin constraint, in addition to Max-V-Place. Candidate (c), with [rˤ], wins because it respects the Split-Margin constraint as well asMax-V-Place and Agree-F-Dors-[rˤ]. This is because the dorsal [u] agrees with the following pharyngealized [rˤ], which also has a dorsal component. The following candidates (d) and (e) lose, because they violate Max-V-Place in addition to Dep and Contig even though (e) respects lowest-ranked Agree-F-Cor. Candidate (f) loses because the pharyngealized [rˤ], which has a dorsal component, is preceded by [i], which is a coronal vowel. The tableau in (12) demonstrates the OT analysis of [u]-epenthesis before the pharyngealized [rˤ] in potential rising-sonority coda clusters for the speakers of the older generation of UHA.

Having provided the analysis for rhotic consonant-to-vowel harmony, the next section discusses and concludes the paper.

4 Discussion and conclusion

This study is one of relatively few studies to discuss sonority- vs. syllable-driven epenthesis in UHA. Other studies do not discuss this in as much detail and it is the only study to provide a detailed OT analysis for the default sonority-driven epenthesis. Furthermore, the study utilized the universal DTE approach in distinguishing between the sonority-driven epenthesis as [i] and syllable-driven epenthesis as [a], based on the prosodic domain of the epenthesis and the quality/sonority of the epenthetic vowel. The result of this is compatible with the universal prediction that [a] is higher in sonority than [i].

The study elaborated on two types of epenthesis in UHA: SylE and default SonE. In UHA, in order to avoid word-internal superheavy syllables after suffixation, the vowel [a] is epenthesized in the process of resyllabification. The vowel [a], which has higher sonority than any other vowel, strengthens the weak degenerate syllable by forming the nucleus of the newly created syllable, which contains the previously unsyllabified consonant. Therefore, the constraint ranking for this type of epenthesis is $^*\Delta\sigma\leq\{i, u\} \gg {}^*\Delta\sigma\leq a$. With regard to default SonE, the high peripheral vowel, [i], which is less in sonority than the low vowel, is epenthesized in the non-head position. The ranking of the DTE constraints is the opposite of that found in syllable-driven epenthesis. In addition, the universal Place of Articulation Hierarchy plays a role in eliminating the high back vowel, [u], from getting epenthesized in the process of default sonority-driven epenthesis.

The default SonE epenthetic vowel in UHA is /i/. However, all vowels, /a/, /i/, and /u/, can be epenthetic to break up a potential rising sonority coda cluster. This is of course due to the harmony between the epenthetic vowel and the neighboring consonants and stem vowels. If the stem contains a high vowel, /i/ or /u/, this vowel spreads its feature [front] or [back] to the sonority-driven epenthetic vowel. This is because the high vowel spreads its feature to the epenthetic vowel to break up such a rising-sonority coda cluster, by inserting [i] in the newly created syllable after the process of epenthesis.

Further, coda consonants play an important role in determining the quality of the vowel in sonority-driven epenthesis. If the potential rising sonority coda cluster ends in a coronal consonant, the epenthetic vowel is the coronal [i] in the process of what is known as coronal consonant-to-vowel harmony.

In a stem that ends with /r/ when the medial consonant is pharyngeal or laryngeal, [a] is epenthesized in the process of pharyngeal harmony. Medial pharyngeal and laryngeal consonants in such words are the trigger of the epenthesis of the low vowel, [a], which has a pharyngeal component, and which itself is the trigger of the pharyngealization of the [$r^ʕ$].

The study also elaborated on the dorsal consonant-to-vowel harmony in words ending with /$r^ʕ$/ for older speakers. /$r^ʕ$/ is the underlying form and is the uncommon realization of the rhotic. It is de-pharyngealized when it is adjacent to a coronal segment. In a potential rising-sonority coda cluster ending with a rhotic, [u] is epenthesized to prevent such fatal coda clusters from surfacing for the speakers of the older generation. The feature [dorsal] for the pharyngealized /rˤ/ dictates the quality of the epenthetic vowel to be [u], matching in the feature [dorsal] between the pharyngealized rhotic and the epenthetic [u]. In this word, [u] is epenthesized, rather than [a], because the secondary [dorsal] feature of the pharyngealized [rˤ] outranks its secondary pharyngeal feature. While this study has provided a phonological account of epenthesis in UHA, more research needs to be done in the future, specifically, to look at the phonetic properties of epenthetic vowels in comparison to lexical vowels in the dialect, similar to Lebanese Arabic (Hall 2024 [this volume]).

The Hasse diagram in Figure 1 summarizes Consonant-to-Vowel Harmony effect on coda clusters in UHA.

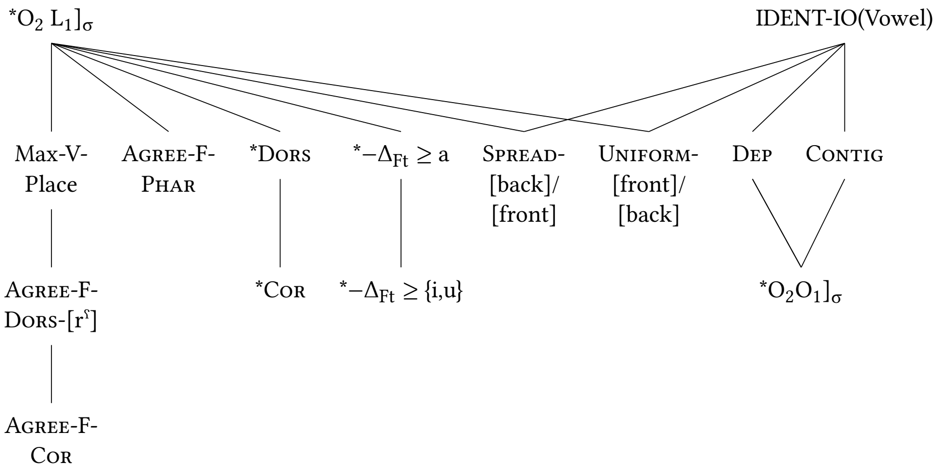


Figure 1: Consonant-to-Vowel Harmony effect on coda clusters in UHA

References

Abu-Mansour, Mahasen. 1987. *A nonlinear analysis of Arabic syllabic phonology, with special reference to Makkan*. University of Florida. (Doctoral dissertation).

Baertsch, Karen. 2002. *An optimality theoretic approach to syllable structure: The split margin hierarchy*. Indiana University Bloomington. (Doctoral dissertation).

Baertsch, Karen & Stuart Davis. 2009. Strength relations between consonants: A syllable-based OT approach. In Kuniya Nasukawa & Phillip Backley (eds.), *Strength relations in phonology*, 285–316. Berlin: Mouton de Gruyter.

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Bokhari, Hassan. 2020. *A comprehensive analysis of coda clusters in Hijazi Arabic: An optimality-theoretic perspective*. Indiana University Bloomington. (Doctoral dissertation).

Broselow, Ellen. 2018. Syllable structure in the dialects of Arabic. In E. Benmamoun & R. Bassiouney (eds.), *The Routledge handbook of Arabic linguistics*, 32–47. London: Routledge.

Clements, Nick. 1991. Place of articulation in consonants and vowels: A unified theory. *Working Papers of the Cornell Phonetics Laboratory* 5. 77–123.

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

Farwaneh, Samira. 2017. Toward a typology of Arabic dialects: The role of final consonantality. *Journal of Arabic and Islamic Studies* 9. 82–109.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Herzallah, Rukayyah S. 1990. *Aspects of palestinian Arabic phonology: A nonlinear approach*. Cornell University. (Doctoral dissertation).

Holt, D. E. 2015. Historical sound change in Optimality Theory. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 545–562. Oxford: Oxford University.

Hume, Elizabeth. 1992. *Front vowels, coronal consonants and their interaction in non-linear phonology*. Cornell University. (Doctoral dissertation).

Hutton, Jeremy. 1996. *Optimality theory and historical language change*. Paper presented at the 4th Phonology Meeting, University of Manchester, U.K.

Kabrah, Rawiah. 2004. *Opacity and transparency in the phonology of Makkan Arabic: A stratal optimality-theoretic approach*. Boston University. (Doctoral dissertation).

Kiparsky, Paul. 2003. Syllables and moras in Arabic. In Caroline Féry & Ruben van de Vijver (eds.), *The syllable in optimality theory*, 147–182. Cambridge: Cambridge University Press.

Lombardi, Linda. 1995. Why place and voice are different: Constraint-specific alternations and optimality theory. *Rutgers Optimality Archive* 105.

Padgett, Jaye. 2011. Consonant-vowel place feature interactions. In Marc van Oostendorp, Colin. J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 3, 1–26. Hoboken: Wiley. DOI: 10.1002/9781444335262.wbctp0075.

Walker, Rachel. 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language & Linguistic Theory* 19. 827–878.

Watson, Janet CE. 2007. Syllabification patterns in Arabic dialects: Long segments and mora sharing. *Phonolgy* 24(2). 335–356.

Younes, Munther. 1993. Emphasis spread in two Arabic dialects. In Mushira Eid & Clive Holes (eds.), *Perspectives on Arabic linguistics V*, 119–145. Amsterdam: John Benjamins.

Youssef, Islam. 2019. The phonology and micro-typology of Arabic. *Glossa: a journal of general linguistics* 4(1). 1–36.

Chapter 5

Segmental and prosodic influences on Bolognese epenthesis

Edward J. Rubin[a] & Aaron Kaplan[a]
[a]University of Utah

Bolognese, the Gallo-Italic grammar of Bologna, eliminates illicit coda clusters via epenthesis. This process is noteworthy for two reasons. First, as in other closely related Romance varieties such as Donceto (Cardinaletti & Repetti 2008), Bolognese prosodic structure and phonotactic patterns determine whether an epenthetic vowel appears: certain clusters are permitted within a PWd but trigger epenthesis when they straddle a PWd boundary, and sonorant-final coda clusters are always subject to epenthesis. Second, Bolognese displays two epenthetic vowels. [u] appears before [v, m], while [e] appears elsewhere. In closely related grammars like Donceto, only one epenthetic vowel ([ə]) appears. We build on Cardinaletti & Repetti's (2008) analysis of coda clusters in Donceto to account for the Bolognese facts.

1 Two contexts for epenthesis

In many ways, pronominal clitics in Bolognese (Romance; Italy) are typical of Romance languages, making the usual morphosyntactic distinctions and mostly exhibiting a typical set of consonants, as highlighted in Table 1. These clitics vary phonologically according to (i) whether they appear as proclitics or enclitics (though the 2p subject clitic (scl) is only enclitic, as in related grammars) and (ii) whether they are adjacent to vowels or consonants. On the other hand, the complexity Bolognese permits in onsets and codas is unusual among Romance languages. Our focus in this paper is these clitics, all of which (in their enclitic

Edward J. Rubin & Aaron Kaplan. 2024. Segmental and prosodic influences on Bolognese epenthesis. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 105–121. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264536

form) display a V ~ ∅ alternation.[1] We argue that this vowel is epenthetic, breaking some clusters up that are permissible in other contexts. In support of this claim we examine other contexts for epenthesis in the language, presenting a unified account for epenthesis in postverbal clitics and these other contexts.

Table 1: Clitic pronouns in Bolognese

	SCL		DCL		ACL		PRT
	Sing	Plur	Sing	Plur	Sing	Plur	
1			**m**	**s**	**m**	**s**	
2	**t**	=**v**	**t**	**v**	**t**	**v**	
3MSG	(a)**l**				(a)**l**		**n**
3RFLX			**s**	**s**	**s**	**s**	

In (1–2) we provide data showing the enclitics that participate in the noted alternation. Like in other Gallo-Italic varieties, subject clitics in Bolognese appear only with tensed verbs and are postverbal in interrogatives (1). The object clitics (dative/indirect (DCL), accusative/direct (ACL), and partitive (PRT)) appear postverbally with tenseless verbs (infinitives, imperatives, gerunds) (2). In both sets of data, we observe the mentioned V ~ ∅ alternation, and the vowels that appear before the relevant consonants, [e] and [u], shall be a main focus of our attention below.

(1) a. i. ˈdoːrm=**e**t — ‘Are you sleeping?’ — 2SG
  ii. durmiˈrɛː=t — ‘Will you sleep?’
 b. i. ˈdoːrm=**e**l — ‘Is he sleeping?’ — 3SG
  ii. durˈme=l — ‘Did he sleep?’
 c. i. _ [2] — 2PL
  ii. durˈmi=v — ‘Are you.PL sleeping?’

[1]The gaps in Table 1 represent non-existent clitics (e.g. reflexive SCLs) or would contain clitics that are or include other vowels and therefore do not participate in this alternation (e.g. the third person dative clitic (DCL) and other third person plural clitics [i], and the third person feminine SCL and all non-reflexive third person accusative clitics (ACLs) which contain [a]). First and second person DCLs and ACLs are both reflexive and non-reflexive. PRT is unspecified for gender and number. All Bolognese data in this paper are drawn from Canepari & Vitali (1995), Vitali (2009), and Lepri & Vitali (2007), or from extensive consultation with native speakers.

[2]There is a predictable gap here: The Vs in question occur only after consonants, and all 2PL tensed verb-forms in Bolognese are V-final. See also footnote 7.

(2)	a.	i.	ˈdɛr=**um**	‘to give me.DCL’	1SG
		ii.	arspuŋˈdiː=m	‘Answer.PL me.DCL!’	
		iii.	truˈvand=**um** iŋ ˈka	‘finding me.ACL at home’	
		iv.	gwaˈrdɛː=m	‘Watch.PL me.ACL!’	
	b.	i.	diˈgaŋd=**et**	‘saying to you.DCL’	2SG
		ii.	ˈda=t	‘Give.SG yourself.DCL … !’	
		iii.	kaˈtɛr=**et**	‘to find/visit you.ACL’	
		iv.	ˈftes=**et**	‘Dress.SG yourself.ACL!’	
	c.	i.	*Non-reflexive* DCL *[i] never alternates this way*		3SG
		ii.	purˈtɛːr=**el**	‘to carry it/him.ACL’	
		iii.	studˈjɛ=l	‘Study.PL it/him.ACL!’	
		iv.	REFL:ˈdɛr=**es**	‘to give to oneself/himself/herself/themselves.DCL’	
		v.	REFL:ˈftaŋd=**es**	‘dressing oneself/himself/herself/themselves.ACL’	
	d.	i.	ˈdɛr=**es**	‘to give us.DCL’	1PL
		ii.	arspuŋˈdiːs	‘Answer.PL us.DCL!’	
		iii.	kaˈtɛr=**es**	‘to find/visit us.ACL’	
		iv.	aspˈtɛː=s	‘Wait.PL for us.ACL!’	
	e.	i.	arspuŋˈdaŋd=**uv**	‘responding to you.PL.DCL’	2PL
		ii.	ˈdɛː=v	‘Give.PL yourselves.DCL … !’	
		iii.	θarˈkɛːr=**uv**	‘to look-for you.PL.ACL’	
		iv.	liˈvɛː=v	‘Get up!/lift yourselves.ACL!’	
	f.	i.	ˈfɛːr=**e**ŋ ˈduː	‘to make two of them’	PRT
		ii.	ʦkuˈræŋn=**e**ŋ ˈdɑp	‘Let’s talk about it later!’	
		iii.	ˈdɛː=ŋ ˈduː a ˈðvaŋ	‘Give.PL two to John!’	

As is apparent, these clitics have the shape [C] following a vowel-final stem, but the shape [eC] following a consonant-final stem (except that we find [um] and [uv], not *[em] and *[ev], for 1SG and 2PL object clitics, respectively; we examine this in Section 2).

Though we will largely ignore preverbal clitics, (3) shows that the [C] form of most of these clitics also appears preverbally (ACLs are shown, and relevant DCLs are identical). In addition, like in many Romance varieties, the preconsonantal ACL.3MSG [al] and ACL.3FSG [la] clitics distinguish gender (3c-i), but this distinction is leveled before a vowel (3c-ii). The corresponding SCL.3SGs [al] / [la] behave identically ([al/la=ˈvad] ‘he/she sees’, [l=e] ‘he/she is’), though the SCL.3MSG has additional allomorphs in preverbal clitic clusters (see Rubin & Kaplan 2022 for an analysis of preverbal clitic allomorphy). The vowel in both of these is distinct from the epenthetic vowels that we discuss below. Note that Bolognese differs importantly from Donceto (another Romance variety spoken near Bolognese in Italy), where the SCL.3MSG includes the epenthetic vowel of that language ([ə]) according to Cardinaletti & Repetti (2008), but the SCL.3FSG includes [a], and undergoes the same pre-V / pre-C allomorphic variation as the two SCL.3SGs in Bolognese. We conclude that Bolognese preverbal SCL.3MSG [al] is due to allomorphy, not epenthesis.

(3)	a.	i.	i=**m**=ˈʦaːmeŋ	‘they call me’	1SG
		ii.	i=**m**=abˈraːθeŋ	‘they hug me’	
	b.	i.	i=**t**=ˈʦaːmeŋ	‘they call you.SG’	2SG
		ii.	i=**t**=abˈraːθeŋ	‘they hug you.SG’	
	c.	i.	i=**al**=ˈʦaːmeŋ	‘they call him’	3SG
			i=**la**=ˈʦaːmeŋ	‘they call her’	
		ii.	i=**l**=abˈraːθeŋ	‘they hug her/him’	
	d.	i.	i=**s**=ˈʦaːmeŋ	‘they call us’	1PL
		ii.	i=**s**=abˈraːθeŋ	‘they hug us’	
	e.	i.	i=**v**=ˈʦaːmeŋ	‘They call you.PL’	2PL
		ii.	i=**v**=abˈraːθeŋ	‘they hug you.PL’	
	f.	i.	i=**i**=ˈʦaːmeŋ	‘they call them’	3PL
		ii.	i=**i**=abˈraːθeŋ	‘they hug them’	
		iii.	i=**s**=ˈʦaːmeŋ	‘they call each other’	
		iv.	i=**s**=abˈraːθeŋ	‘they hug each other’	
	g.	i.	i=**ŋ**=ˈʦaːmeŋ ˈduː	‘they call two of them’	PRT
		ii.	i=**n**=abˈraːθeŋ ˈduː	‘they hug two of them’	

The choice between the [C] and [VC] forms of the enclitics in (1–2) is dictated by Bolognese’s coda cluster phonotactics. Two phonotactic requirements are relevant: a prohibition on coda clusters in certain prosodic domains and a prohibition on sonorant-final coda clusters. We begin with the former.

A variety of coda clusters is attested in the language; this includes clusters ending with [s] or [t] as we see in (4). Interestingly, though, these clusters are not permitted when the final [s] or [t] is a clitic, as (5) shows. The bolded epenthetic vowels in (5) break up the clusters in these examples, and in some examples it is the minimal difference with a correspondent in (4).

(4) [Cs] and [Ct] can occur word-finally...

a.	skɛːrs	‘scarce’
b.	sɛːlt	‘(a) jump’
c.	ˈt=sɛːlt	‘you jump’
d.	a=ˈpæŋs	‘I think’
e.	a=trˈavers	‘I cross’
f.	peːrs	‘lost’
g.	t=iŋˈvæŋt	‘you invent’
h.	a=gˈwaːst	‘I spoil’

(5) ...but epenthesis occurs when the [s] or [t] is a clitic

a.	ˈskɛːr=**e**s	‘to dry us/ourselves’
b.	ˈsɛːl=**e**t	‘do you salt (something)?’
c.	amiˈrɛːr=**e**s	‘to admire us/ourselves’
d.	liˈvɛːr=**e**s	‘to get us/ourselves up’
e.	liˈvɛːr=**e**t	‘to get you/yourself up’
f.	truˈvɛːr=**e**s	‘to find us/ourselves’
g.	truˈvɛːr=**e**t	‘to find you/yourself’
h.	ˈrɑŋf=**e**t	‘do you snore?’

The contrast between (4) and (5) indicates that enclitics and the verb are (immediate) constituents of distinct prosodic units that differ in whether they permit epenthesis. Cardinaletti & Repetti (2008) document similar facts for subject clitics in Donceto, but they do not provide an explicit analysis. They argue that clitics are outside the prosodic word (PWd) but within the phonological phrase (PP), and we adopt that position here.[3] To illustrate, the structure of [ˈsɛːl=**e**t] ‘do you salt (something)?’ is given in Figure 1, setting aside the epenthetic vowel.

Such an analysis leads to the following generalization: a C + [s]/[t] coda is permitted PWd-internally, but when it straddles a PWd boundary it is banned.

[3]For present purposes, the identities of the relevant prosodic categories are unimportant. They might be PWd and PP, or perhaps recursive PWds (Itô & Mester 2007). We adopt the former to follow Cardinaletti & Repetti’s precedent.

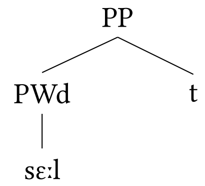


Figure 1: Prosodic structure of [ˈsɛːl=et] ‘do you salt (something)?’

To account for this, we posit that *Complex outranks Contiguity(PP) but not Contiguity(PWd).[4] In Table 2, the /lt/ cluster is contained within the root and is thus PWd-internal; Contiguity(PWd) blocks epenthesis (because the PWd is a constituent of the PP, the cluster is also PP-internal, hence candidate (b)’s Contiguity(PP) violation). But in Table 3, the cluster is not wholly within the PWd (because the /t/ is a clitic) and is therefore subject only to the low-ranking Contiguity(PP); this time, *Complex compels epenthesis.[5]

Table 2: /ˈsɛːlt/ ‘(a) jump’, from (4)

/ˈsɛːlt/	Contig(PWd)	*Complex	Contig(PP)
☞ a. ˈsɛːlt		*	
b. ˈsɛːlet	*!		*

Evidence that the epenthetic vowels in these forms are indeed epenthetic comes from two sources. First, as we have seen, these clitics do not always surface with [e] (6), appearing as just [s] or [t] when doing so does not violate *Complex. Furthermore, some object clitics, including the ones at issue here, have a [VC] allomorph that appears, for example, after the second singular subject clitic, but the vowel that appears in this allomorph is [a], not [e] (7).

[4]Onset clusters behave somewhat differently, suggesting a distinction between *ComplexOnset and *ComplexCoda. Because we will not analyze onset clusters here, we will simply use *Complex.

[5]Contiguity(PP) and Contiguity(PWd) are in a stringent relationship (De Lacy 2004): assuming PWds are always, or at least usually, contained within PPs (whether one adopts the strict layer hypothesis (e.g. Selkirk 1984b) or something else), any configuration subject to Contiguity(PWd) is also subject to Contiguity(PP). A prediction of this analysis is therefore that whatever the ranking between these two constraints, if epenthesis or any other Contiguity-violating process is blocked in elements outside a PWd but within a PP, it will also be blocked inside a PWd. But the opposite does not hold: as in Bolognese, epenthesis may occur within a PP even if it is blocked within a PWd.

Table 3: /ˈsɛːlet/ ‘Do you salt (something)?’, from (5), Figure 1

/ˈsɛːl=t/	Contig(PWd)	*Complex	Contig(PP)
a. ˈsɛːlt		*!	
☞ b. ˈsɛːlet			*

(6) a. i=**s**=ˈsakeŋ ‘they dry us’
b. al=**s**=aˈmiːra ‘he admires us’
c. a=**s**=iŋdurmiŋˈtæŋ ‘we fall asleep’
d. a=**t**=ˈtroːv ‘I find you’
e. ˈ**t**=sɛːlt ‘you jump’
f. ˈ**t**=rɑŋf ‘you snore’

(7) a. t=**at**=iŋdurˈmæŋt ‘you fall asleep’
b. t=**as**=ˈtroːv ‘you find us’

Second, epenthesis in contexts not involving clitics uses the same vowel that we see in (5). For example, despite the ranking Contiguity(PWd) ≫ *Complex, PWd-internal epenthesis to break up coda clusters is attested; some examples are given in (8) (we address the cause of this epenthesis below). Each of these roots contains a root-final cluster. In the first form on a line, a suffix allows the second of those consonants to surface as an onset, avoiding a coda cluster. But in the second form on a line, in the absence of suffixes, [e] is epenthesized between the consonants. Aside from regular exceptions to be discussed in Section 2, the vowel that appears in these contexts is always [e].

(8) a. fsg [-a] / fpl [-∅]
i. ˈtɛːvla / ˈtɛːv**e**l ‘table’ / ‘tables’
ii. laŋˈteːrna / laŋˈteːr**e**ŋ ‘lantern’ / ‘lanterns’
iii. ˈliːvra /ˈliːv**e**r ‘hare’ / ‘hares’
b. infinitive [-ˈɛr] / pres.1sg [-∅]
i. sfitˈlɛːr / a=ˈsfat**e**l ‘to slice’ / ‘I slice’
ii. urdˈnɛːr / a=ˈɑʊrd**e**ŋ ‘to order’ / ‘I order’
iii. lusˈtrɛːr / a=ˈlost**e**r ‘to polish’ / ‘I polish’

c. ADJ$_{FSG}$ [-a] / ADJ$_{MSG}$ [-∅]

i.	ˈdabla / ˈdab**e**l	‘weak.FS’ / ‘weak.MS’
ii.	ˈðɑʊvna / ˈðɑʊv**e**ŋ	‘young.FS’ / ‘young.MS’
iii.	ˈvɔːstra / ˈvɔːst**e**r	‘your.FS’ / ‘your.MS’

The evidence therefore suggests that the [e] seen in (5) is epenthetic. This conclusion ties the appearance of this vowel to other patterns of epenthesis in Bolognese, and it is simpler than an alternative that posits two [VC] allomorphs for these clitics, one with [e] that appears only word-finally and one with [a] that appears elsewhere.

More must be said about (8). We attribute the epenthesis illustrated there to phonotactic requirements. It is tempting to view that epenthesis as a manifestation of sonority sequencing principles (see Selkirk 1984a and Clements 1990 for overviews) that prohibit rising-sonority coda clusters (e.g. *[a=ˈsfatl], *[a=ɑʊrdŋ]) or clusters that do not have an adequate fall in sonority (*[laŋˈteːrŋ], plausibly). But it is actually unclear to what extent Bolognese obeys sonority sequencing constraints. A representative sample of the language’s coda clusters is given in Table 4; see also (4). Most clusters conform to sonority sequencing expectations, but not all do (e.g. [rbz], [dg]); onset clusters are even more dramatic in their disregard for sonority sequencing ([zbdɛl] ‘hospital’, [ˈftleŋna] ‘slice’, [ˈʦkɲɔser] ‘to disavow’, [vdand] ‘seeing’). One clear generalization, though, is that sonorant-final coda clusters are unattested, and we therefore adopt a constraint against such clusters, *C[+SON]]$_{\sigma}$, and this constraint drives epenthesis in (8).

Table 4: Licit obstruent coda clusters

rbz	fo**rbz**	‘scissors’	rð	zgɛː**rð**	‘wool comb’
rb	tɑʊ**rb**	‘cloudy’	dg	ˈapɛː**dg**	‘I walk’
rp	auˈzuː**rp**	‘I usurp’	mb	strapˈjɑ**mb**	‘overhang’
rd	sɑʊ**rd**	‘deaf’	mg	ˈstɑ**mg**	‘stomach’
rdg	poː**rdg**	‘portico’	mt	ˈgɑ**mt**	‘elbow’
rt	pɛː**rt**	‘part’	ŋdg	pɑ**ŋdg**	‘mouse’
rʦ	kwɛː**rʦ**	‘lid’	ŋf	gre**ŋf**	‘claw’
rθ	pɔː**rθ**	‘pig’	ŋp	ka**ŋp**	‘field’

*C[+SON]]$_{\sigma}$ outranks CONTIG(PWd), as illustrated in Table 5.

To summarize, we have identified two considerations that drive epenthesis in Bolognese. The first is *COMPLEX, whose effect is visible outside the PWd, trigger-

Table 5: ˈliːver ‘hares’, from (8)

/ˈliːvr-∅/	*C[+son]]$_\sigma$	Contig(PWd)	*Complex	Contig(PP)
a. ˈliːvr	*!		*	
☞ b. ˈliːver		*		*

ing epenthesis in final clusters involving consonantal enclitics. Within the PWd, epenthesis eradicates sonorant-final coda clusters.

So far we have dealt only with examples in which the epenthetic vowel is [e], but in certain situations [u] appears instead. We turn now to those contexts.

2 Epenthetic [u]

As we have said, the primary epenthetic vowel in Bolognese is [e], which surfaces in a variety of contexts. But when followed by a labial consonant, the epenthetic vowel is instead [u]. For example, [m] – being a sonorant – unsurprisingly triggers epenthesis in coda clusters, just like the other sonorants shown in (8). But the epenthetic vowel that precedes [m] is [u]:

(9) a. fsg [-a] / fpl [-∅]
   i. ˈaːnma / ˈaːn**u**m ‘soul’ / ‘souls’
   ii. ˈfɑʊrma / ˈfɑʊr**u**m ‘form’ / ‘forms’
   iii. baˈtaɪzma / baˈtaɪz**u**m ‘baptism’ / ‘baptisms’

b. infinitive [-ˈɛr] / pres.1sg [-∅]
   i. kalˈmɛːr / a=ˈkɛːl**u**m ‘to calm’ / ‘I calm’
   ii. lagarˈmɛːr / a=ˈlɛːgr**u**m ‘to weep’ / ‘I weep’
   iii. farˈmɛːr / a=ˈfaɪr**u**m ‘to stop’ / ‘I stop’

c. adj$_{\text{fsg}}$ [-a] / adj$_{\text{msg}}$ [-∅]
   i. ˈuːltma / ˈuːlt**u**m ‘last.fs’ / ‘last.ms’
   ii. ˈsɛːtma / ˈsɛːt**u**m ‘seventh.fs’ / ‘seventh.ms’

We attribute the appearance of [u] in (9) to Agree(lab)-rime (10) (see, e.g., Lombardi 1999 for discussion of Agree constraints). Under the assumption that the distinction between round and unround vowels is captured formally by the feature [labial] (as opposed to [round]; Clements 1991), this constraint can compel epenthesis of a round vowel like [u] when the following coda consonant is

[+labial]. AGREE(lab)-rime holds only for segments appearing in the same rime; evidence for this restriction on the constraint's effect is presented below.

(10) AGREE(lab)-rime: within a rime, adjacent segments must match for [labial].

Epenthesis of [u] is an example of The Emergence of the Unmarked (McCarthy & Prince 1994). AGREE(lab)-rime is outranked by IDENT(labial), which prevents underlying vowels from becoming round to match a following labial coda. As (11) shows, vowel quality before [m] is not generally restricted. But epenthetic vowels have no input correspondent, and AGREE(lab)-rime can influence their realization.

(11)
a. θ**i**mˈzɛːra 'bedbug infestation'
b. ʤ**e**mˈleŋ 'gem.DIM'
c. pr**e**m 'first'
d. krizaŋˈt**e**ːm 'chrysanthemum'
e. ˈ**ɑ**mbra 'shadow'
f. eˈkɔn**o**m 'treasurer'
g. ˈ**o**md 'humid'

The effect of AGREE(lab)-rime is illustrated in Table 6. [6] *C[+SON]]$_\sigma$ compels epenthesis, and AGREE(lab)-rime selects the candidate with an epenthetic [u].

Table 6: ˈaːn**um** 'souls', from (9)

/ˈaːnm-∅/	*C[+SON]]$_\sigma$	IDENT(lab)	AGR(lab)-rime	CONTIG(PWd)	*COMPLEX	CONTIG (PP)	*[V, +lab]	*[V, +hi]
a. ˈaːnm	*!				*			
b. ˈaːnem			*!	*		*		
☞ c. ˈaːnum				*		*	*	*

[6] To keep the tableau simple, *COMPLEX and the CONTIGUITY constraints are omitted from most subsequent tableaux in this section. As Table 6 shows, they are ranked too low to affect the outcome in the kinds of cases presently under consideration.

Epenthesis of [e] when the relevant coda consonant is not [+labial] has two possible explanations. Either AGREE(lab)-rime requires the epenthetic vowel to match the coda consonant's [−labial] specification, or AGREE(lab)-rime is ambivalent in the face of a [−labial] coda consonant and the constraints *[V, +lab] and *[V, +hi] favor [e]. For purposes of illustration, we adopt the former approach.

Curiously, labial consonants trigger the appearance of the high vowel [u], not a mid vowel [o] or [ø], either of which would satisfy *[V, +hi] and be more similar to the default [e]. To account for this we adopt *RoLo and *RoFro (Archangeli & Pulleyblank 1994, Kaun 1995, 2004), which prohibit round non-high vowels and round front vowels, respectively. Their effect is visible in Table 7: candidate (a) is eliminated by *C[+SON]]$_\sigma$, and of the remaining candidates, only candidate (f) satisfies AGREE(lab)-rime, *RoLo, and *RoFro.

Table 7: ˈaːn**um** 'souls', from (9)

/ˈaːnm-∅/	*C[+SON]]$_\sigma$	IDENT(lab)	AGR(lab)-rime	*RoLo	*RoFro	*[V, +lab]	*[V, +hi]
a. ˈaːnm	*!						
b. ˈaːnem			*!				
c. ˈaːnom				*!		*	
d. ˈaːnøm				*!	*!	*	
e. ˈaːnym					*!	*	*
☞ f. ˈaːnum						*	*

To our knowledge, the only other labial consonant that triggers a preceding epenthetic vowel is [v] (cf. forms with coda-cluster-final [b, p, f] in Table 4). As our analysis predicts, that epenthetic vowel is [u]:

(12) a. FSG [-a] / FPL [-∅]
- i. ˈseːrva / ˈseː**ruv** 'servant' / 'servants'
- ii. ˈkaˈteːrva / kˈateː**ruv** 'multitude' / 'multitudes'

b. INFINITIVE [-ˈɛr] / PRES.1SG [-∅]
- i. kurˈvɛːr / a=ˈkuː**ruv** 'to bend' / 'I bend'
- ii. userˈvɛːr / t=u'seː**ruv** 'to observe' / 'I observe'

c. N_{FSG} [-a] / N_{MSG} [-∅]
   i. ˈvadva /ˈvad**u**v — ‘widow’ / ‘widower’
d. N_{MSG}.DIM [-ˈeŋ] / N_{MSG} [-∅]
   i. narˈveŋ /ˈneːr**u**v — ‘little nerve’ / ‘nerve’

As with [m], Agree(lab)-rime favors [u] to match the [+labial] [v]. What is most notable about these examples, however, is that [v] triggers epenthesis in the first place. As shown in (4) and Table 4, epenthesis does not usually occur when a cluster ends with an obstruent. We argue in Section 3 that [v] is in fact a sonorant in Bolognese, so the examples in (12) simply further illustrate epenthesis in sonorant-final clusters. Before turning to that argument, however, some loose ends need attention.

As discussed above, Ident(labial) prevents Agree(lab)-rime from causing non-epenthetic vowels to change. This is illustrated in Table 8 with the form [iŋˈkɛːv] ‘groove.’ The underlying /ɛː/ surfaces faithfully despite the following [v].

Table 8: iŋˈkɛːv ‘groove’

/iŋˈkɛːv/	*C[+son]]$_\sigma$	Ident(lab)	Agr(lab)-rime	*RoLo	*RoFro	*[V, +rnd]	*[V, +hi]
☞ a. iŋˈkɛːv			*				*
b. iŋˈkuːv		*!				*	**

Furthermore, examples such as [ˈliːv=**et**] ‘Get up!’ (with an epenthetic [e]) show that labial consonants trigger rounding only on preceding epenthetic vowels, not following ones. The difference, we suggest, is that when an epenthetic vowel precedes a labial (or any) consonant, that consonant is invariably a coda – hence Agree(lab)-rime’s requirement that only segments in the same rime must match for labiality.

The evaluation of [ˈliːv=**et**] is shown in Table 9. *Complex compels epenthesis here because the form contains an enclitic, and this time Agree(lab)-rime favors an unround vowel, eliminating candidate (c); recall that if Agree(lab)-rime is interpreted to be ambivalent in this case, lower constraints favor [e], too.

Finally, the data in (13) show the convergence of the two environments for epenthesis that we have focused on here. These examples show epenthesis in

Table 9: ˈliːv=**et** 'Get up!'

/ˈliːv=t/	*C[+son]]$_\sigma$	Ident(lab)	Agr(lab)-rime	*RoLo	*RoFro	*[V, +rnd]	*[V, +hi]
a. ˈliːv=t	*!						*
☞ b. ˈliːv=et							*
c. ˈliːv=ut			*!			*	**

word-final clusters involving clitics, driven by *Complex. Here, though, the clitics are sonorants and are therefore subject to *C[+son]]$_\sigma$. Epenthesis occurs exactly as expected: [u] appears before [m] and [v], and [e] appears elsewhere. A representative tableau is shown in Table 10.

(13) a. ˈliːv=**el** 'Is he lifting (something) up?' / 'lift him up!'
b. ˈliːv=**e**ŋ 'lift some up!'
c. ˈliːv=**u**m 'lift me up!'
d. liˈvɛːr=**u**v 'to lift you up'

Table 10: ˈliːv=**el** 'Is he lifting (something) up?' / 'lift him up!'

/ˈliːv=l/	*C[+son]]$_\sigma$	Contig(PWd)	*Complex	Contig(PP)
a. ˈliːvl	*!		*	
☞ b. ˈliːvel				*

3 The status of [v]

Padgett (2002) argues that in Russian and possibly other languages, the segment transcribed as [v] is more properly treated as a sonorant. Bolognese appears to belong to this group of languages. The contrast between, on the one hand, (12), with [v]-final clusters, and (8) and (9), with sonorant-final clusters, and, on the other hand, (4)/Table 4, with clusters with other obstruents in final position (including [b, p, f]), is just one piece of evidence for this position.

In addition, Bolognese [v] sometimes alternates with [w], as in (14).

(14) a. ˈakwa / ˈakuv ‘water’ / ‘waters’
b. iŋsiŋˈwɛr / t=iŋˈsiːnuv ‘to insinuate’ / ‘you insinuate’
c. koŋˈtiːgwa / koŋˈtiːguv ‘contiguous.FS’ / ‘contiguous.MS’

Moreover, Canepari & Vitali (1995: 148) write:

> /v/ often vanishes: [farˈa(v)ɑŋna] ‘guinea fowl’, [(v)ɲ o] ‘come (past part.)’ (or also [farˈaʋɑŋna]); occasionally it becomes [w]: [asˈwad] ‘si vede/one sees’.

Some sources (including Canepari & Vitali in the excerpt just provided) transcribe this sound as [ʋ], indicating that some listeners hear the sound as an approximant, not a fricative.[7]

4 Discussion and conclusion

Like other Romance languages, Bolognese shows epenthesis that is sensitive to morphological and prosodic structure. Our account of this extends the analysis of Cardinaletti & Repetti (2008) to account for a collection of facts that are peculiar to Bolognese, such as variation in the quality of the epenthetic vowel, the avoidance of sonorant-final coda clusters, and [v]’s patterning with sonorants.

Chief among our claims is that the alternation between [C] and [VC] seen in Bolognese’s enclitics involves epenthesis rather than deletion or suppletion. This position has two major benefits. First, it connects clitic allomorphy to broader epenthetic processes in the language. Second, these clitics show extensive allomorphy (Rubin & Kaplan 2022), and treating some of this allomorphy as epenthesis reduces the number of allomorphs in the lexicon and/or the number of clitic-specific processes that must be posited.

The analysis presented here represents just a first attempt to account for the interaction between vowel epenthesis and Bolognese’s clitics. Our focus has been on word-final clusters, but proclitics are also subject to epenthesis in familiar ways: [e] is epenthesized, except that [u] appears when sharing a rime with [m] or [v]:

[7] Historically, Bolognese [∅/v/ʋ/w] comes from Latin [w]; perhaps it has not (yet?) fully transitioned from a sonorant to an obstruent. Perhaps relatedly, across all conjugations, in the imperfect and conditional, the stress is penultimate, and not final as with all other 2PL verb-forms. With those two forms, in the interrogative, no enclitic 2PL SCL [v] is present, e.g. [maɲaˈresi=∅] ‘Would you.PL eat?’ With all other 2PL verb-forms, with final stress, an enclitic [v] does appear, whether SCL, DCL, or ACL. Adjacency to primary stress plays a clear role, perhaps both diachronically and synchronically.

(15) a. al=v**e**=ˈdɛːva ‘He was waking you up.’
b. al=m**e**=ˈʦftes ‘He’s undressing me.’
c. l=**u**m=ˈda ‘he gives me’
d. i=s**e**=ˈfteːveŋ ‘they were getting dressed’
e. a=s**e**=ˈfteːveŋ ‘we were getting dressed’

Notice that [e], not [u], appears before [f] in the final two examples in (15). Several explanations are available: perhaps [f] is syllabified as an onset here ([ft] clusters are attested in Bolognese); perhaps the Agree constraint used above might be further restricted to *sonorants* that share a rime; or perhaps [e] is not epenthetic here, but rather part of the underlying representation of the reflexive clitic [se] and other relevant clitics. Furthermore, alongside the similarities between proclitics and enclitics are substantive differences. In particular Rubin & Kaplan (2022) argue that proclitics exhibit rather extensive allomorphy that enclitics do not. One example, of many similar, concerns the 3msg subject and direct object proclitics (as noted in the discussion above (3)), which regularly exhibit a form [al] that never occurs as an enclitic ([t=al=ˈvad]/*[t=el=ˈvad] ‘you see him’ vs. [ˈvadr=el]/*[ˈvadr=al] ‘to see him’) even though as an enclitic [al] would satisfy *C[+son]]$_\sigma$ just as well as [e] epenthesis does. Another issue arises when both the subject and object clitics are 3ms and adjacent, which can only occur preverbally. The [al] allomorph is not permitted here for both, and epenthesis arises, though after the object clitic, rather than before it ([al=le=ˈvad]/*[al=al=ˈvad]/*[al=el=ˈvad] ‘he sees him’).

Our supposition is that at least some of these differences arise from differences in combinatorial possibilities, as with the consecutive 3ms clitics, which occur only preverbally. As another example, the 3ms.nom proclitic appears as [l] before vowels; when 3ms.nom is an enclitic it is always verb-final and therefore never prevocalic in isolation. But in context, when the following word begins with a vowel, [l] appears.

In sum, the analysis developed here supplies a foundation on which a broader treatment of clitic alternations in Bolognese, and indeed perhaps those found in other languages, can be built.

Acknowledgements

The data in this paper were mostly collected or checked in fieldwork over several years with prominent members of the native speaker community that centers in the cultural association “Al Pånt dla Biånnda.” Data were also drawn from

Canepari & Vitali (1995), Vitali (2009), and Lepri & Vitali (2007). We thank our consultants for their patient help with the data, and we thank Roberto Serra and Daniele Vitali for extensive discussion of it. Thanks also to the audience at *Epenthesis and Beyond* for feedback on this work.

References

Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded phonology*. Cambridge: MIT Press.

Canepari, Luciano & Daniele Vitali. 1995. Pronuncia e grafia del Bolognese. *Rivista Italiana di Dialettologia* 19. 119–164.

Cardinaletti, Anna & Lori Repetti. 2008. The phonology and syntax of preverbal and postverbal subject clitics in Northern Italian dialects. *Linguistic Inquiry* 39(4). 523–563.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In John Kingston & Mary Beckman (eds.), *Between the grammar and physics of speech* (Papers in Laboratory Phonology 1), 283–333. New York: Cambridge University Press.

Clements, Nick. 1991. Place of articulation in consonants and vowels: A unified theory. 5. 677–123.

De Lacy, Paul V. 2004. Markedness conflation in Optimality Theory. *Phonology* 21(2). 145–199.

Itô, Junko & Armin Mester. 2007. Prosodic adjunction in Japanese compounds. In Yoichi Miyamoto & Masao Ochi (eds.), *Formal approaches to Japanese linguistics: Proceedings of FAJL 4. MIT Working Papers in Linguistics*, 97–111. Cambridge: MIT Department of Linguistics & Philosophy.

Kaun, Abigail. 1995. *The typology of rounding harmony: An Optimality Theoretic approach*. Los Angeles: UCLA. (Doctoral dissertation).

Kaun, Abigail. 2004. The typology of rounding harmony. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, 87–116. Cambridge: Cambridge University Press.

Lepri, Luigi & Daniele Vitali. 2007. *Dizionario bolognese-italiano italiano-bolognese*. Bologna: Pendragon.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in optimality theory. *Natural Language & Linguistic Theory* 17(2). 267–302.

McCarthy, John J. & Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistic Society* 24. 18.

Padgett, Jaye. 2002. *Russian voicing assimilation, final devoicing, and the problem of [v].* Ms., UC Santa Cruz.

Rubin, Edward J. & Aaron Kaplan. 2022. Lexical selection in Bolognese clitic allomorphy. *Isogloss. Open Journal of Romance Linguistics* 8(5). 1–17. DOI: 10.5565/rev/isogloss.223.

Selkirk, Elisabeth O. 1984a. On the major class features and syllable theory. In Mark Aronoff & Richard T. Oehrle (eds.), *Language sound structure: Studies in Phonology*, 107–136. Cambridge: MIT Press.

Selkirk, Elisabeth O. 1984b. *Phonology and syntax: The relation between sound and structure.* Cambridge: The MIT Press.

Vitali, Daniele. 2009. *Dscårret in bulgnaiṡ?* 2nd edn. Bologna: Perdisa.

Chapter 6

Epenthesis as a matter of FAITH

Christian Uffmann
HHU Düsseldorf

In Optimality Theory, the observation that default epenthetic segments are sourced from a heavily restricted set of segments across languages, such as the glottal stop [ʔ] for consonants, schwa [ə] or [i] for vowels, is generally analysed as a markedness effect: Optimal epenthetic segments are maximally unmarked segments. This chapter highlights several conceptual and empirical problems with this claim and proposes a faithfulness-based alternative instead: Optimal epenthetic segments involve minimal epenthesis at the level of distinctive features, that is, minimal violations of Dep(F). We will show how this proposal accounts for crosslinguistically common epenthetic segments as segments that are in some way underspecified vis-à-vis other segments, and it will be shown how the problems encountered in a markedness-based approach can be resolved.

1 Introduction

Crosslinguistically, the set of possible default segments in epenthesis seems quite heavily restricted. Typical default consonants are glottals, especially the glottal stop [ʔ], while typical vowels are schwa [ə] or [i]. This raises the question of what principles restrict the set of possible epenthetic segments, or put differently: what makes [ʔ] or [ə] a good epenthetic segment?[1]

Before we continue, a few brief words of clarification. When I discuss default epenthesis in this chapter, I mean the phonologically motivated insertion of a contextually invariant segment. By contextual invariance I mean that segment

[1]There is the open question to what extent there are also marked or 'unnatural' epenthetic segments (see e.g. Vaux & Samuels 2017). In this chapter I will leave the question open but simply note that the claim is contentious.

Christian Uffmann. 2024. Epenthesis as a matter of FAITH. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 123–141. Berlin: Language Science Press.
DOI: 10.5281/zenodo.14264538

quality is not influenced by adjacent segments (as in vowel copy or glide insertion). In addition, epenthesis has to be phonologically motivated, that is, epenthesis is prosodically or phonotactically optimising. This includes cases of consonant epenthesis to resolve hiatus or to satisfy an onset requirement, or cases of vowel epenthesis to break up illicit consonant clusters or to satisfy coda requirements. Crucially, this definition excludes cases of epenthesis that are either morphologically conditioned or morphologically restricted (such as cases of epenthesis that only occur with certain affixes).[2]

As an example of phonologically motivated epenthesis of a default segment, consider glottal stop epenthesis in German. A glottal stop is inserted (a) in hiatus position, before a stressed syllable, (b) word-initially, and (c) stem-initially (see the examples in (1).

(1) Glottal stop epenthesis in German (Alber 2001, Wiese 1998)

a.	[ʁuˈʔiːn]	‘ruin’	[koˈʔaːlɐ]	‘koala’
b.	[ʔɛlç]	‘moose’	[ʔoˈʔaːzə]	‘oasis’
c.	[ˈʔapˌʔaʁtɪç]	‘deviant’	[ˈmɪtˌʔaʁbaɪt]	‘cooperation’

So why is the glottal stop inserted in German (and many other languages)? In Optimality Theory (Prince & Smolensky 1993, henceforth OT), the standard answer has been that default epenthesis constitutes a case of the emergence of the unmarked: optimal epenthetic segments are universally unmarked segments (Lombardi 2002, 2003, Lacy 2006). In this chapter, I want to propose a different account of default epenthesis, one that is based on faithfulness instead, more precisely feature faithfulness. Optimal epenthetic segments are not unmarked; instead they insert as few distinctive features as possible.

The next section will briefly introduce the markedness approach to epenthesis and raise a few questions and problems with respect to this approach. §3 will introduce the alternative, the faithfulness-based account and discuss how it addresses the questions and problems raised by the markedness approach. In §4 I will discuss how the faithfulness-based approach fits in with other approaches to epenthesis and conclude.

2 Epenthesis of the unmarked?

In Optimality Theory, default segment epenthesis is commonly analysed as a markedness effect, thereby also explaining why the set of possible epenthetic

[2] For an overview of other types of epenthesis and a discussion of phonological and morphological epenthesis, the reader is referred to Żygis (2010).

consonants seems to be quite heavily restricted (Lombardi 2002, 2003, Lacy 2006). Default segment epenthesis is epenthesis of the least marked segment, following universal markedness hierarchies (Lombardi 2002, 2003) or markedness scales (Lacy 2006). De Lacy is adamant that it can only be markedness that is responsible for the selection of the optimal epenthetic segment, as the other main constraint family, faithfulness, cannot play a role in this selection: All segments violate the segmental anti-insertion constraint DEP while vacuously satisfying the feature faithfulness constraint IDENT – since no features in the inserted segment have an input correspondent. Let us now look at how this model accounts for the limited set of segments found as default epenthetic segments, that is glottal segments, especially [ʔ], in the case of consonants and [ə, i] in the case of vowels.

To account for glottals as default epenthetic consonants, Lombardi (2002) and Lacy (2006) propose to extend the established place markedness hierarchy, according to which [coronal] is less marked than [labial] and [dorsal] by adding [pharyngeal] (Lombardi) or [glottal] (de Lacy) at the bottom of the hierarchy. Glottals are thus the least marked consonants in terms of place of articulation and therefore selected for epenthesis.

This approach also makes a second prediction, namely that in cases where glottals are unavailable for independent reasons, coronal epenthesis will be found instead, and it is here that we find a first potential wrinkle with the markedness approach to epenthesis: Coronal default epenthesis is only marginally attested, if at all. The textbook case of [t]-epenthesis in Axininca Campa has been convincingly reanalysed as deletion in the inverse context (Staroverov 2015), and the other examples Lombardi (2002) mentions also generally do not qualify as instances of purely phonologically driven default epenthesis or are amenable to different analyses, by Lombardi's own admission. The markedness approach thus overpredicts variation in consonant epenthesis.

Quite the opposite situation holds when we look at default epenthetic vowels, where considerable variation is found. Lombardi (2003) nevertheless attempts to provide a markedness-based account of this variation. She argues that [ɨ] is the least marked vowel, followed by [ə] and then [i, a], and if one vowel is unavailable in a language, the next one will be selected. This finding again translates into universal markedness scales. The unmarkedness of the central vowels is formalised by stating that [+back, −round] vowels are least marked (central vowels being analysed as back unrounded), and the preference of [i, ɨ] is analysed as the preference of high vowels over mid vowels. Yet, despite allowing variation, some attested default vowels are not predicted by Lombardi's model, such as [e], found for example in Spanish and Gengbe (Archangeli 1988, Abaglo & Archangeli 1989) or Bolognese (Rubin & Kaplan 2024 [this volume]). Lacy

(2006) concedes that markedness only has a limited role to play in the selection of default epenthetic vowels (mostly by excluding rounded vowels) and suggests sonority as an additional factor, with different languages preferring high-, low- or mid-sonority vowels. We thus see the opposite situation for vowels than for consonants: Markedness alone underpredicts the range of attested epenthetic vowels, and this raises the question of where this asymmetry between consonants and vowels comes from: Why is there considerably more crosslinguistic variation regarding default vowels compared to consonants?

It also raises a more general question: What evidence is there, outside epenthesis, that default segments actually are the crosslinguistically least marked segments?[3] The concept of markedness is, of course, complex (see e.g. Rice 2007 for an overview), and a detailed discussion would go beyond the scope of this chapter, but let us use crosslinguistic frequency as a proxy for evaluating the markedness of a segment. Frequency is a well-motivated diagnostic especially in the OT framework, as we should expect segments to appear in many languages when the constraints against them are universally very low ranked (given that there are no conditions on inputs, a principle known as Richness of the Base, and phoneme systems follow epiphenomenally from constraint rankings). I thus used the phoneme database PHOIBLE (Moran et al. 2014) to establish how crosslinguistically common default epenthetic segments are in phoneme systems (Table 1).

Table 1: Crosslinguistic frequencies of default segments in PHOIBLE

segment	in % of languages	rank
[ʔ]	37	19
[ɨ]	16	17
[ə]	22	11
[i]	99	1

We can see that the glottal stop is crosslinguistically not particularly frequent, found in little more than a third of languages, being only the 19th most frequent consonant, a somewhat disappointing showing for what is supposedly the least marked consonant of all.

[3]Note also that Lombardi has to classify pharyngeals as the least marked consonants and back unrounded vowels as the least marked vowels, which is especially surprising, as the general assumption is that rounding is unmarked on back vowels, and that back vowels are more marked than front vowels.

The results for vowels are even more perplexing. The allegedly least marked vowel, [ɨ], is rather infrequent, found in only 16% of all languages surveyed, and schwa only fares marginally better. [i] aligns well with the markedness approach, though: It is the most frequent of all vowels and practically universally present in vowel inventories. Otherwise, frequencies are the opposite of what the markedness scale should lead us to expect, the least marked vowel (according to Lombardi) actually being the least frequent.

It is not just crosslinguistic frequency, however. In languages that have glottal segments or schwa, these are often restricted in their distribution. In English and German, for example, schwa is found in unstressed syllables only, while the phonemic glottal /h/ is restricted to word- or foot-initial singleton onsets. Such contextual restrictions are also surprising for allegedly unmarked segments.

Moreover, in some languages these segments are in fact restricted to epenthetic contexts. There are languages in which a default epenthetic segment is not otherwise found in the phoneme inventory, for example [ʔ] in German (Alber 2001, Wiese 1998) or schwa in Italian (Repetti 2012), Western Aramaic (Eid & Plag 2021) and Anindilyakwa (Mansfield et al. 2024 [this volume]). As Krämer (2006) first noticed, this creates a ranking paradox: In order to be selected as the optimal epenthetic segment, the constraints against that segment have to be ranked low(est), while in order to be excluded elsewhere (if posited as underlying), the same constraints need to rank relatively high in order to prevent the segment from surfacing faithfully. So in German the fact that there is no phonemic glottal stop would standardly be analysed by assuming a high-ranked constraint *[ʔ]. In order to be epenthetic, however, *[ʔ] needs to rank below all other segmental (consonantal) markedness constraints.[4] The fact that the default epenthetic segment may not be permissible elsewhere in a language thus poses a serious problem for the markedness-based approach to epenthesis.

There is one final issue with the markedness-based approach I want to discuss here. Default epenthesis is not the only available epenthesis strategy. Instead, epenthetic segment quality can also be determined, fully or partially, by spreading or copying features from adjacent segments.[5] There are cases of vowel copy

[4]Krämer's analysis relies on Comparative Markedness. Alternatively, one could invoke positional markedness to prohibit the segment from all contexts except those where it happens to be epenthetic. Both approaches look unsatisfying, though. While Krämer has to rely on essentially arbitrary rerankings of what should be universal markedness scales, the alternative approach has to stipulate that the epenthetic segment is only permissible in epenthesis contexts via a brute force mechanism.

[5]I will use the term 'spreading' without necessarily implying autosegmental spreading but using it as a theory-neutral term instead.

and vowel harmony (see e.g. Stanton & Zukoff 2018), there is spreading from vowels to consonants, as in glide insertion, where the glide typically agrees in backness with a neighbouring vowel, and there are also cases where consonants spread features to vowels, as in labial attraction, where vowels are rounded next to consonants (as in Bolognese; Rubin & Kaplan 2024 [this volume]).

The markedness-based approach cannot account for these alternative types of epenthesis, and additional mechanisms have to be introduced to handle these, without there being a connection between the two types of epenthesis or a principled explanation when which kind is found. That such a connection exists is clear, however, from cases where the different types of epenthesis interact. Uffmann (2007, 2006) shows that in loanword adaptation, all three processes – vowel harmony, consonantal spreading, and default epenthesis – are frequently found. This is not limited to loanword adaptation, however. A case in point is Urban Hijazi Arabic, where we find all three strategies in vowel epenthesis into final clusters of rising sonority (for a detailed discussion, see Bokhari 2024 [this volume]). The relevant data from Al-Mohanna (2021) are given in (2); generalisations are my own.

(2) Epenthesis in Urban Hijazi Arabic (Al-Mohanna 2021)

a.	ʔisim	/ism/	‘name’	ʃukur	/ʃukr/	‘gratitude’
b.	ʔakil	/akl/	‘food’	madiħ	/madħ/	‘praise’
c.	tamur	/tamr/	‘dates’	baħar	/baħr/	‘sea’

The data show that vowel copy is the most common strategy, but it is restricted to the high vowels [i, u] (examples in (a)). When the stem vowel is /a/, however (Urban Hijazi Arabic has the common Arabic 3-vowel system), as in (b), we generally find epenthesis of default [i] (also found in other epenthesis contexts), unless the epenthetic vowel is preceded by a labial or pharyngeal consonant; in these cases we find spreading of the consonantal place feature, yielding [u] after labials and [a] after pharyngeals (c). How can such a conspiracy of epenthesis strategies be analysed insightfully, if default epenthesis and copy or spreading epenthesis constitute separate phenomena?

To summarise the foregoing, the markedness approach into epenthesis runs into a number of empirical and theoretical problems, listed here again:

- There is an asymmetry between consonants and vowels: the markedness approach predicts more variation regarding consonant epenthesis than is attested, but cannot account for the full range of variation attested for default vowels.

- Typologically common epenthetic segments are probably not unmarked; they are crosslinguistically not very frequent and often positionally restricted in languages that have them.
- The default vowel or consonant in a language may otherwise be absent from that language's phoneme inventory.
- There is no unified approach to default and copy epenthesis.

The next section will therefore introduce an alternative proposal, arguing that default epenthesis is not based on markedness but on faithfulness instead, more precisely on feature faithfulness. I will propose that optimal epenthetic segments minimally insert features, and I will then show how this approach can account for the problems raised above.

3 Epenthesis of the faithful!

The alternative proposal I am going to outline in this section is that epenthesis in general minimally violates the feature faithfulness constraint DEP(F) (Zoll 1996): when segments are inserted, a minimal amount of features is inserted along with the segment. Default segment epenthesis then is not a case of the emergence of the unmarked but a case of the emergence of the unspecified or underspecified. This, of course, presupposes that segments differ in their featural complexity, so that less complex segments are better candidates for epenthesis. I will motivate this assumption in a moment, but let us first take a brief historical detour, as the idea that epenthesis is connected to underspecification is not new.

This idea was explicitly pursued in underspecification theory, most notably radical underspecification (Archangeli 1988, Pulleyblank 1988). In this theory, segments are specified minimally; predictable feature values are inserted by rule in the course of the derivation. As a consequence, there is one vowel and one consonant in every language that carries no feature specifications in the underlying representations but receives them via rule instead. When a segment is inserted, only the segmental slot is inserted, and then the feature fill-in rules that exist independently will take care of the segmental content of the epenthetic segment. As a result, the default epenthetic segment of a language is always the segment that is underlyingly unspecified. For example, in a language like Spanish where the default vowel is [e], underlying /e/ is unspecified, and then there is a set of rules that fills in default feature specifications, in Spanish [-high, -back, -round], yielding surface [e], when no specifications are present. Differences in default segments between languages stem from differences in these fill-in rules.

This general approach is, of course, not available in OT. There are no serial derivations in the course of which feature specifications can be inserted, and there are no constraints on inputs that could, for example, prohibit fully specified input segments. The general idea that epenthesis involves underspecified segments can be transferred to an OT approach, however, as I will argue now.

In OT, segments can also show different degrees of feature specification, in three ways. First, there is considerable evidence that some features at least are privative, not binary (see e.g. Clements & Hume 1995, Lombardi 1996, Anderson & Ewen 1987, Beckman et al. 2013). Privative features are also a prime reason to invoke a class of Dep(F) constraints in addition to (or replacing) Ident constraints, which operate on binary features. Now if there are indeed privative features, different degrees of segmental complexity follow naturally: segments are or are not specified for some feature; some segments will consequently carry very few features.

Second, while radical underspecification with feature fill-in is not an option in OT, the same cannot be said about theories of contrastive underspecification, such as Dresher (2009), especially if we assume that underspecified segments remain surface-underspecified (no fill-in; the interface to phonetics interprets underspecified segments, as in Hall 2011). Consequently, not all segments are specified for all features. Features that are redundant for a class of segments (say, [voice] for sonorants) may be absent in the feature specifications of this class, again yielding different degrees of segmental complexity. Feature privativity and contrastive underspecification can also be combined as in Iosad (2012).

Even if readers remain skeptical of these additional assumptions (and motivating them in detail would go beyond the scope of this chapter), there is a third way in which segments can show different degrees of specification or complexity, one that does not need additional and perhaps controversial assumptions about the nature of phonological representations. Some segments are intrinsically less specified than others. It is commonly assumed that glottal segments are not specified for oral features (or do not have an Oral / Supralaryngeal node in feature geometry), thus making them intrinsically less complex than oral segments, while schwa is often analysed as a featureless segment (e.g. van Oostendorp 2000, Crosswhite 2004). I want to argue that this is the true reason for them being good default segments, rather than their unmarkedness. I will now develop this argument and show how it also addresses the questions and problems with the markedness approach discussed above.

For consonants, glottals are selected for epenthesis not because they are the least marked segments (they may in fact be fairly marked), but because they are intrinsically less complex than oral segments, lacking oral features. Consequently, they violate the constraint against feature insertion DEP(F) less than oral segments (no insertion of place and manner features). With this proposal we can also address and explain the ranking paradox mentioned above, that the default epenthetic segment in a language may otherwise be absent from the phoneme inventory of that language. Let us illustrate this with a toy language example, modelled on the German facts mentioned earlier.

To begin with, our toy language has no phonemic glottal stop. Hypothetical underlying glottal stops will be deleted.[6] In OT this is easily modelled by ranking a constraint against glottal stops *ʔ above the no-deletion constraint MAX (see Table 2).

Table 2: Glottal stop deletion...

/aʔ/	*ʔ	MAX
a. [aʔ]	*!	
☞ b. [a]		*

Now assume our toy language also inserts glottal stops to satisfy an onset requirement. We will thus add ONSET to the tableau as a trigger for epenthesis. Now what drives the selection of the glottal stop as an optimal epenthetic consonant is the constraint DEP(F) ranked above *ʔ, as in (3).

Table 3: ...and epenthesis

/aʔ/	ONSET	DEP(F)	*ʔ	MAX	DEP	*t
a. [aʔ]	*!		*			
b. [a]	*!			*		
c. [ʔaʔ]		{Lar}	**!		*	
☞ d. [ʔa]		{Lar}	*	*	*	
e. [ta]		{Lar, Oral!}		*	*	*

The first two candidates replicate (3) but do not satisfy ONSET and are therefore eliminated. The last two candidates are relevant, [ʔa] vs. [ta]. Although *ʔ outranks *t – assuming that [t] is part of the language's segment inventory –

[6]As the alternative, buccalisation, i.e change into an oral segment, is robustly unattested typologically, it is not considered here.

[ʔ] is selected for epenthesis because it fares better on Dep(F): it only incurs violations for the insertion of laryngeal features, while [t]-epenthesis also incurs violations for inserting oral features. Finally, the third candidate [ʔaʔ] shows that *ʔ is still relevant: even when a glottal stop is inserted, underlying glottal stops will still be deleted; the second violation of *ʔ proves fatal for this candidate. In sum, the feature faithfulness approach to epenthesis can explain why a segment can be epenthetic even though it is not in a language's phoneme inventory (and thus relatively marked in that language).

A closer look at the proposed analysis reveals that it does not preclude markedness-driven epenthesis: With Dep(F) sufficiently low-ranked, markedness constraints can still determine the quality of the epenthetic segment. While this is a theoretical possibility, such a ranking is very unlikely to ever arise in practice, however. To see why, consider briefly the consequences of low-ranked Dep(F), beyond epenthesis. It would mean that all segments are changed whenever adding a feature decreases their markedness. This is an across-the-board change, and it will thus trigger Lexicon Optimisation: the altered segments will be posited as underlying (because in the absence of alternations learners have no evidence to the contrary), but this in turn removes evidence for low-ranked Dep(F). In short, even if there is a stage where Dep(F) is ranked low, the changes this induces on underlying forms will ensure that the constraint is quickly promoted. It is therefore highly unlikely that there could ever be a stable period of low-ranked Dep(F) that would be necessary for a markedness-based epenthesis process to emerge.

But let us turn to vowels now. Recall that vowels display considerably greater variation when it comes to the quality of the default epenthetic vowel, another challenge for the markedness-based approach. How can we explain this? I propose that this is because there is no class of vowels that is inherently less specified than others, unlike the class of glottals in the set of consonants, with one exception: There is schwa, which is featureless; hence schwa epenthesis does not incur any Dep(F) violations, making it an ideal candidate for epenthesis. The reason that we do not find schwa in many languages can be motivated by a constraint against featureless vowels (such as *[] in McCarthy 2018). Once schwa is out of the race, there is no obvious alternative candidate that is underspecified. Instead, underspecification and markedness may both play a role in selecting the optimal epenthetic vowel.

Both [i] and [e] are attested epenthetic vowels. While the insertion of [i] can be explained as a markedness effect, the insertion of [e] cannot, as mid vowels are considered more marked than both high and low vowels (see e.g. Lombardi 2003). However, [e]-insertion could be explained as an underspecification effect. If we assume that features are privative and that [i] is [high], then [e] is underspecified for height (neither [high] nor [low]), and this favours its insertion. Let us look at a little toy language again to see how epenthesis of [e] can be modelled as an underspecification effect.

In our toy language, vowel (see the tableau in (4)) epenthesis is triggered by a high-ranked NoCoda constraint. Consider three candidate vowels: [i], [e], [ə]. Assume that both [i] and [e] are specified as [coronal] (front) vowels, and [i] is additionally specified as [high], while [ə] is featureless. While [i] is less marked than [e] (the ranking *Mid » *High, see Lombardi 2003), [e] is nevertheless the more harmonic candidate vowel, as it violates Dep(F) less. Schwa satisfies Dep(F) but is ruled out by a high-ranked constraint against featureless segments, *[].

Table 4: Default epenthesis of [e]

/pat/	NoCoda	*[]	Dep(F)	*Mid	*High
a. [pat]	*!				
b. [pati]			{Cor, Hi!}		*
☞ c. [pate]			{Cor}	*	
d. [patə]		*!			

It is not clear to what extent markedness is necessary at all to explain the variation between [i] and [e] in epenthesis. Recall an older (and still ongoing) debate in phonology whether vowel height is marked by the feature [high] or by [open] (leaving high vowels underspecified), with arguments going both ways (see e.g. Clements 1990). The choice of aperture feature may hence well be language-specific with evidence coming from the phonological behaviour of the vowels. Are high vowels or mid vowels phonologically active in a language? The underspecification approach to epenthesis makes an interesting prediction here (though one that goes well beyond the remit of this chapter): that languages where [high] is active should prefer a mid vowel for epenthesis, and vice versa. Romance languages may be a case in point: High vowels trigger metaphony processes in many varieties, and mid vowels are generally found epenthetically. At any rate, language-dependent specifications (dependent on the phonological activity of vowels, see also Dresher 2009 for discussion) may also yield different default epenthetic vowels. Further research should be able to shed light on this question and the potential link between feature activity and default epenthesis.

There is one more point we need to address: the connection between default epenthesis and copy or spreading-based epenthesis that a markedness-based approach cannot provide. Faithfulness provides such a connection, as both default epenthesis and spreading are different ways of minimising Dep(F) violations, assuming that spreading does not involve the insertion of a new feature but just the extension of its domain, for example by adding an association line.

Assume further that spreading comes at a cost, violating anti-spreading constraints. The interaction we find between spreading-based and default epenthesis then follows from the relative ranking of anti-spreading constraints with Dep(F). Uffmann (2006, 2007) suggests families of anti-spreading constraints, also taking into account non-local spreading and feature sharing between consonants and vowels. For the present purpose, assume a simple constraint *Share.

(3) *Share(F)
Every feature F is associated with exactly one segment

Let us briefly return to the Urban Hijazi data discussed earlier. High vowels spread to an epenthetic vowel slot (/i, u/ are copied) while there is epenthesis of the default vowel [i] after low /a/. This indicates that [high] spreads more easily than [low]. We can therefore posit a ranking *Share(lo) » *Share(hi) with Dep(F) sandwiched in between: Spreading high vowels is more harmonic than inserting a feature, which in turn is more harmonic than low vowel spreading.[7] The tableaux in (5-6) illustrate this. For each input form there are three candidates: a faithful candidate with no epenthesis violating the Sonority Sequencing Principle (SSP), one candidate where the epenthetic vowel results from spreading (indicated by the tie bar between the vowels), and one candidate with default [i] insertion.

Table 5: High vowel spreading...

/ʃukr/ 'gratitude'	SSP	*Share(lo)	Dep(hi)	*Share(hi)
a. [ʃukr]	*!			
☞ b. [ʃukur]				*
c. [ʃukir]			*!	

Epenthesis is triggered by the high-ranked SSP constraint, eliminating the fully faithful candidate. In (5) spreading is optimal vis-a-vis default epenthesis because *Share(hi) ranks below Dep(F). In (6) default epenthesis is optimal because Dep(F) ranks below *Share(lo).

[7]This raises the question to what extent the proposed ranking *Share(lo) » *Share(hi) is language-specific or universal. A detailed discussion of this question would go beyond the scope of this chapter, but note that Uffmann (2007) proposes universal spreading hierarchies, derived from established markedness hierarchies.

Table 6: ...but default epenthesis after low vowels

/ʔakl/ ‘food’	SSP	*Share(lo)	Dep(hi)	*Share(hi)
a. [ʔakl]	*!			
b. [ʔakal]		*!		
☞ c. [ʔakil]			*	

Note that this approach makes an interesting prediction: that such interactions between default epenthesis and spreading should be restricted to languages where the epenthetic vowel is not schwa. As schwa epenthesis satisfies both Dep(F) and *Share, it should always be optimal vis-a-vis spreading. A preliminary survey seems to confirm this prediction, but further research could shed more light on this question.

To summarise, I proposed that the selection of the default epenthetic segment is not determined by markedness but by faithfulness, more precisely satisfaction of Dep(F), the constraint against feature insertion. Under this view, glottals are optimal epenthetic consonants because they are inherently featurally less complex, not having any oral segments. For default vowels more variation is expected as there is no vowel, bar featureless schwa, that is inherently less specified than others. As epenthesis is uncoupled from markedness, this can also explain why there are languages where the default segment is not part of the phoneme inventory. Finally, this approach can unify default epenthesis and spreading or copy epenthesis, as both result from the same drive, to minimise Dep(F) violations.

4 Discussion and conclusions

To conclude this chapter, I want to cast the net wider and discuss briefly how the faithfulness-based account of epenthesis relates to other approaches to epenthesis, and I want to argue that it fits in well with observations made in these approaches. I want to discuss three approaches briefly: gradual epenthesis in serial OT, perceptually-based epenthesis and the relationship between epenthetic and excrescent or intrusive segments, with a look at Articulatory Phonology.

In serial versions of OT, the candidate generator function Gen can only perform one change to the input at a time. The winning candidate is then iteratively resubmitted to evaluation until no more improvements are possible. This raises the question of what a single change is, and McCarthy (2008, 2018) argues that segment deletions are not a single change, but the end point of a series of deletions of features or feature-geometric nodes. He furthermore proposes that consonant deletion always goes through a stage of debuccalisation before full deletion. Epenthesis should therefore also be gradual, a point taken up by Al-Mohanna (2021). Now if epenthesis is the mirror image of deletion, and debuccalisation is the final step before deletion, as argued by McCarthy, then adding laryngeal specifications should be the first step in epenthesis, and after that it should be hard or even impossible to add any additional features and still remain harmonically optimising, as buccalisation is hardly attested in the world's languages,[8] while debuccalisation is common (and improves harmony, to allow for deletions). In other words, once laryngeal specifications are inserted, further harmonic improvement is unlikely and the derivation stops. Glottal epenthesis is therefore predicted as a consequence of gradual feature insertion. For vowels, no comparable 'brake' exists, and feature additions are possible as long as they are harmonically optimising, but they will still be minimal: Serial OT predicts that the featurally least specified possible surface segment is selected and thus supports the faithfulness-based approach to epenthesis.

Another approach to epenthesis that has received particular attention in loanword adaptation is perceptually based. The idea is that a segment is inserted that is perceptually closest to zero. In other words, the epenthetic segment should be as imperceptible as possible (see e.g. Kenstowicz 2007). Steriade (2009) introduces the idea of the P-map to this effect, a map of segment confusabilities. These are translated into perceptibility scales, which can then be implemented as scalar constraints. Steriade proposes scalar faithfulness constraints: the more perceptible a segment is, the less likely is its insertion or deletion. Hence, for example, the optimality of schwa epenthesis would derive from the lower perceptibility of schwa compared to other vowels, formalised as scalar Dep-IO constraints: Dep(i) » Dep(ə). The P-map approach thus also appeals to faithfulness, but differently, by using faithfulness constraints as anchors for the perceptibility scales rather than by appealing to faithfulness per se.

I would argue instead that the insights from the perceptual approach to epenthesis can be accounted for in a faithfulness-based approach without having to appeal to additional sets of constraints. Assuming – uncontroversially, I hope – that features are at the interface to phonetic pairings of articulatory gestures with salient acoustic cues,[9] the fewer features a segment has, the fewer salient cues its articulation contains, and the less generally perceptible (or confusable with

[8]The only convincing case I can think of is in Northern Pame (Berthiaume 2003: 226), where a sequence of two glottal segments turns into a glottalised lateral.

[9]For a formalisation see e.g. Boersma & Hamann (2009) within the framework of Bidirectional Phonology.

zero) it will be. It may thus well be possible to subsume the perceptual approach under my proposal.

Finally, there is an ongoing and renewed debate regarding the distinction between intrusive or excrescent segments and true phonological epenthesis (see e.g. Hall 2006, 2024 [this volume], Krämer 2024 [this volume], Bellik 2024 [this volume]), which can sometimes be difficult to distinguish. Besides, epenthesis can diachronically result from phonologising intrusion (Hall 2006, Karlin 2021), which is commonly analysed as gesture retiming, especially gestural underlap, leading to the perception of an additional segment although no articulatory gestures are added. When intrusion phonologises, gesture addition should remain minimal, and thus feature insertion should remain minimal (as features correspond to gestures on the phonological level). This is in fact the strong view of Articulatory Phonology (Browman & Goldstein 1992): no gestures may be added to an underlying form (Gick 1999), that is, all epenthetic segments must make use of underlyingly present gestures, and all that happens in a phonological derivation is the retiming and reorganisation of an underlying gestural score. While this claim may after all be too strong, it too ties in with the idea pursued in this chapter, that epenthesis involves the minimal insertion of features. The markedness-based approach, in contrast, has nothing to say about the relationship between epenthesis and intrusion; why should the phonologisation of gestural underlap result in a universally unmarked segment?

To conclude, I argued in this chapter that the standard markedness-based approach to epenthesis in Optimality Theory is beset with theoretical and empirical problems, and I sketched, however briefly, an alternative approach that is based on minimal feature insertion instead. This was formally captured in OT by the DEP(F) constraint, addressing and resolving the problems that the markedness-based approach encounters, and tying in with independent work on epenthesis. Future research should test the validity of this approach further.

Acknowledgements

I would like to thank audiences at the Epenthesis Workshop and at the University of Düsseldorf and two anonymous reviewers for discussion and feedback. All remaining errors are my own.

References

Abaglo, Poovi & Diana Archangeli. 1989. Language-specific underspecification: Gengbe /e/ and Yoruba /i/. *Linguistic Inquiry* 20. 457–480.

Alber, Birgit. 2001. Regional variation and edges: Glottal stop epenthesis and dissimilation in standard and southern varieties of German. *Zeitschrift für Sprachwissenschaft* 20. 3–41.

Anderson, John & Colin Ewen. 1987. *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Archangeli, Diana. 1988. Aspects of underspecification theory. *Phonology* 5(2). 183–207.

Beckman, Jill, Michael Jessen & Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. True voice languages. *Journal of Linguistics* 49. 259–284.

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Berthiaume, Scott. 2003. *A phonological grammar of Northern Pame*. University of Texas at Arlington. (Doctoral dissertation).

Boersma, Paul & Silke Hamann. 2009. Loanword adaptation as first-language phonological perception. In Andrea Calabrese & Leo Wetzels (eds.), *Loan phonology*, 11–58. Amsterdam: Benjamins.

Bokhari, Hassan. 2024. The patterning of epenthesis in Urban Hijazi Arabic. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 85–104. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264534.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3-4). 155–180.

Clements, George N. 1990. The hierarchical representation of vowel height. Ms. Cornell University.

Clements, George N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, 245–306. Oxford: Blackwell.

Crosswhite, Catherine. 2004. Vowel reduction. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetic bases of markedness*, 191–231. Cambridge: Cambridge University Press.

Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Eid, Ghattas & Ingo Plag. 2021. Syllable structure and syllabification in Maaloula Aramaic. Ms. HHU Düsseldorf.

Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16. 29–54.

Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28(01). 1–54.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Iosad, Pavel. 2012. *Representation and variation in substance-free phonology: A case study in Celtic*. Universitetet i Tromsø. (Doctoral dissertation).

Karlin, Robin. 2021. Finnish inserted vowels: A case of phonologized excrescence. *Nordic Journal of Linguistics* 45. 49–79.

Kenstowicz, Michael. 2007. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences* 29. Ms. MIT, 316–340.

Krämer, Martin. 2006. The emergence of the comparatively unmarked. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 236–244. Somerville: Cascadilla Proceedings Project.

Krämer, Martin. 2024. Prokaryotic syllables and excrescent vowels in two Yuman languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 225–245. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264546.

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

Lombardi, Linda. 1996. Postlexical rules and the status of privative features. *Phonology* 13. 1–38.

Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19. 219–251.

Lombardi, Linda. 2003. Markedness and the typology of epenthetic vowels. Ms. University of Maryland.

Mansfield, John, Rosey Billington & Hywel Stoakes. 2024. Vowel predictability and omission in Anindilyakwa. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 57–84. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264532.

McCarthy, John. 2008. The gradual path to cluster simplification. *Phonology* 25. 271–319.

McCarthy, John. 2018. How to delete. In Amel Khalfaoui & Matthew Tucker (eds.), *Perspectives on Arabic linguistics XXX*, 7–32. Amsterdam: Benjamins.

Al-Mohanna, Faisal. 2021. Gradual vowel epenthesis in Urban Hijazi Arabic. *Languages* 6(160). 15 pp.

Moran, Steven, Daniel McCloy & Richard Wright. 2014. *PHOIBLE online*. Tech. rep. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms. Rutgers University and University of Colorado at Boulder.

Pulleyblank, Douglas. 1988. Vocalic underspecification in Yoruba. *Linguistic Inquiry* 19. 233–270.

Repetti, Lori. 2012. Consonant-final loanwords and epenthetic vowels in Italian. *Catalan Journal of Linguistics* 11. 167–188.

Rice, Keren. 2007. Markedness in phonology. In Paul V. De Lacy (ed.), *The Cambridge handbook of phonology*, 79–98. Cambridge: Cambridge University Press.

Rubin, Edward J. & Aaron Kaplan. 2024. Segmental and prosodic influences on Bolognese epenthesis. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 105–121. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264536.

Stanton, Juliet & Sam Zukoff. 2018. Prosodic identity in copy epenthesis: Evidence for a correspondence-based approach. *NLLT* 36. 637–684.

Staroverov, Peter. 2015. [t]-epenthesis revisited. The case of Apurucayali. Ms. Universität Leipzig.

Steriade, Donca. 2009. The phonology of perceptibility effects: The P-Map and its consequences for constraint organization. In Kristin Hanson & Sharon Inkelas (eds.), *The nature of the word: Studies in honor of Paul Kiparsky*, 151–79. Cambridge: MIT Press.

Uffmann, Christian. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116. 1079–1111.

Uffmann, Christian. 2007. *Vowel epenthesis in loanword adaptation*. Tübingen: Niemeyer.

van Oostendorp, Marc. 2000. *Phonological projection: A theory of feature content and prosodic structure*. Berlin: de Gruyter.

Vaux, Bert & Bridget Samuels. 2017. Consonant epenthesis and markedness. In Bridget Samuels (ed.), *Beyond markedness in formal phonology*, 69–100. Amsterdam: Benjamins.

Wiese, Richard. 1998. *The phonology of German*. Oxford: Clarendon.

Zoll, Cheryl. 1996. *Parsing below the segment in a constraint-based framework*. UC Berkeley. (Doctoral dissertation).

Żygis, Marzena. 2010. Typology of consonantal insertions. *ZAS Papers in Linguistics* 52. 111–140.

Chapter 7

Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study

Jennifer Bellik
UC Santa Cruz

Onset clusters in Turkish loanwords have previously been described as being repaired with an optional epenthetic vowel, as in Turkish *spor* [sɯpor] 'sport' (Clements & Sezer 1982, *inter alia*). However, the percept of an inserted vowel can also result from gestural timing (Gafos 2002, Hall 2003). Bellik (2019b) presents acoustic evidence that the vocoids in Turkish onset clusters are gradiently present and differ acoustically from their underlyingly present counterparts, arguing that they arise from gestural timing, not epenthesis. This chapter presents ultrasound data from five Turkish speakers that further support this interpretation. An SSANOVA analysis shows that the tongue body is in a significantly different position during the interconsonantal interval (ICI) in underlying #CC words compared to in underlying vowels in #CVC words. Anticipatory coarticulation with the following vowel appears to influence tongue position in the ICI in #CC more than in #CVC. I argue that the process is best understood as vowel intrusion combined with coarticulation.

1 Introduction

Turkish phonology is well-known for its highly regular vowel harmony system. Canonically, all non-initial vowels in a word match the preceding vowel in backness, and high vowels additionally match the preceding vowel in rounding. Synchronically, this harmony applies to vowels in (most) suffixes: *ip-ler-in-iz* 'your (pl.) ropes' but *pul-lar-ın-ız* 'your (pl.) stamps'. Vowel harmony has not

Jennifer Bellik. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540

applied within loanwords, however, so Turkish includes many words where vowels disagree in backness, rounding, or both, such as *kitap* 'book' (Arabic), *broşür* 'brochure' (French), and *İstanbul* (Greek). When a word contains disharmonic vowels, harmonizing vowels match the immediately preceding vowel: *kitap-lar-ınız* 'your (pl.) books', *gid-iyor-uz* 'we are going'. Turkish is typically described as permitting some complex codas but prohibiting all complex onsets (Clements & Sezer 1982); both types of consonant cluster occur only in loanwards. Vowel insertion occurs in complex codas with rising sonority and in any onset cluster. These two vowel insertion processes differ in several ways. Vowel insertion in coda clusters (Table 1) is obligatory, written, and invariant. The inserted vowel can receive primary stress, and harmonizes in backness and rounding with the preceding lexical vowel, like other high vowels in Turkish suffixes. Coda clusters generally occur in words of Arabic or Farsi origin.

Table 1: Vowel insertion in Turkish coda clusters

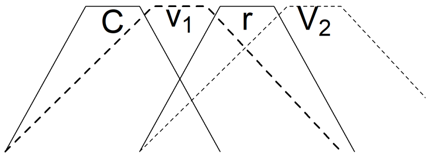
Spelling	Nom. (epenthesis, no suffix)	Acc. (-ɯ/i/u/y)	Dat. (-e/a)	Root	Gloss
sabır	[sa.bɯr]	[sab.ˈrɯ]	[sab.ˈra]	/sabr/	'patience'
cebir	[dʒe.ˈbir]	[dʒeb.ˈri]	[dʒeb.ˈre]	/dʒebr/	'algebra'
burun	[bu.ˈrun]	[bur.ˈnu]	[bur.ˈna]	/burn/	'nose'
ömür	[ø.ˈmyr]	[øm.ˈry]	[øm.ˈre]	/ømr/	'life'

Vowel insertion also occurs in onset clusters, which usually appear in borrowings from European languages (Table 2; Yavaş 1980, Clements & Sezer 1982, Kaun 2000, Yıldız 2010, Kabak 2011). These inserted vowels are variable and unwritten, and have been characterized as optional (Yıldız 2010) or style-dependent (Clements & Sezer 1982). They sometimes match the following lexical vowel in backness and/or rounding, but are always [+back] following /g/ or /k/ (Clements & Sezer 1982, Kabak 2011).

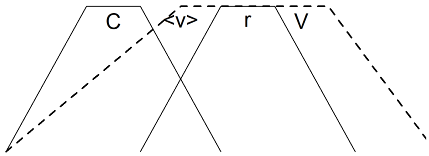
Though both types of insertion have previously been characterized as epenthesis, this unified description does not explain the differences between them. These differences would be explained, however, if the presence of vowels in codas is due to actual vowel insertion (that is, epenthesis driven by syllable structure), while vowels in onsets are due only to gestural alignment, as will be illustrated in Figure 1, which I have argued previously (Bellik 2018, 2019a,b).

Table 2: Vowel insertion in Turkish onset clusters

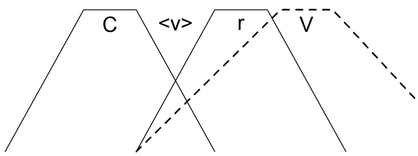
Spelling	With insertion	Without insertion	Gloss
branda	[buranda]	[branda]	‘canvas’
prens	[pirens]	[prens]	‘prince’
prova	[purova]	[prova]	‘test’
blujin	[buluʒin] ~[byluʒin]	[bluʒin]	‘blue jeans’
grip	[gurip]	[grip]	‘flu’



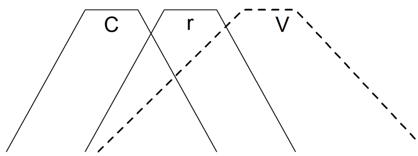
(a) Epenthesis (V gesture inserted between C and r)



(b) Copy-vowel intrusion (Open C-r transition, more V-ICI overlap)



(c) Schwa-like intrusion (Open C-r transition, less V-ICI overlap)



(d) Transparent complex onset (Close C-r transition)

Figure 1: Four possible gestural scores for /CrV/

Figure 1 shows four possible gestural scores that could be the output of phonology, given /CrV/ as the input, ranging from epenthesis in Figure 1a to a canonical complex onset in Figure 1d. In a canonical complex onset, the first consonant’s gesture is immediately followed by the second consonant’s. No vowel sound heard between them, because /r/ achieves its closure before /C/’s release. In vowel intrusion (Figure 1b,c), the two consonant gestures are likewise adjacent, with no distinct vowel gesture intervening. Vowel intrusion results when the first consonant’s closure is released before the second consonant achieves its closure, opening the vocal tract during the transition between consonants. This open transition allows an overlapping vowel gesture to be heard before the second consonant recloses the vocal tract (Browman & Goldstein 1993, Gafos 2002, Hall 2003, 2006). Intrusion can be realized in a variety of ways, depending on the precise gestural alignments involved. If /V/ achieves its target during the interconsonantal interval (ICI), and the ICI is long enough, the resulting open transition can sound like a copy of the following /V/ (Figure 1b). When the ICI is short, and/or /V/ does not attain its target during the ICI (Figure 1c), the open transition sounds more schwa-like, and may be more affected by the preceding consonant than a longer "copy" vowel. Since inserted vowels in Turkish onset clusters reportedly take on the backness, but not the height, of V2 (Yavaş 1980, Clements & Sezer 1982, Kaun 2000, Yıldız 2010, Kabak 2011), they can be considered intermediate between schwa-like vowels and copy vowels.

Intrusion resembles epenthesis (Figure 1a) in that the first consonant's closure is released before the second consonant's closure is achieved, and in that acoustically, a period of high amplitude periodicity intervenes between consonants. They differ, however, in that in epenthesis, an additional vowel gesture intervenes between the consonants. Consequently, tongue position in the ICI reflects a phonologically specified tongue body target. This is also true for an underlying /CVC/ sequence. By contrast, in intrusion, the acoustic "vowel" between the consonants is only a preview of the following vowel, audible during the open transition between consonants. Therefore, tongue position during the ICI reflects the transition between the preceding consonant and the following vowel. Unlike an epenthetic vowel, the intrusive vowel is not a phonological segment, and therefore not a target for any segment-level vowel harmony.

The phonological status of vowels in Turkish onset clusters has implications for our understanding of Turkish syllable structure, vowel harmony, and their interaction. Turkish may offer an example of vowel intrusion in a language that also has a phonological process of vowel harmony, broadening the range of descriptions of vowel harmony processes. Morever, the phonological status of these inserted vowels is of broader theoretical importance, because a variety of claims about the Turkish or cross-linguistic vowel harmony have been founded on the interpretation that inserted vowels in Turkish onset clusters are epenthetic targets for vowel harmony (Clements & Sezer 1982, Yavaş 1980, Kaun 2000, Yıldız 2010).

I hypothesize that vowels in Turkish onset clusters are intrusive, resulting from gestural timing, and therefore differ gesturally and acoustically from underlying vowels. I tested this hypothesis with an ultrasound and acoustic production study. The acoustic study (Bellik 2018, 2019a) found, firstly, that the duration of the ICI in /CC/ words has a unimodal distribution, indicating that the vowels heard in onset clusters are gradiently present. If they were optional epenthetic vowels with a specified durational target, they would be either categorically present (one mode at a positive vocalic duration) or categorically absent (another mode at zero vocalic duration), creating a bimodal distribution of durations instead. Secondly, the F1 and F2 values of vowels in onset clusters before /i/ and /o/ differ from both harmonizing vowels /i/ and /u/ (V2 determines which vowel is harmonic) and the non-harmonizing vowel /ɯ/. If vowels in onset clusters are epenthetic and had acoustic targets, we would expect them to hit the same targets that underlyingly present vowels hit: either the same F1~F2 values as harmonizing vowels in the same context, if vowel harmony applies; or the same F1~F2 values as the lexical /ɯ/, which is the Turkish phoneme that these vowels, impressionistically, most resemble, as well as being the most schwa-like Turkish vowel, suggesting that it could function as a neutral vowel. But in fact, vowels in onset clusters differ acoustically from both the phonemes harmony predicts and the phoneme they most sound like, suggesting that vowels in onset clusters do not share the acoustic targets of any lexical vowels. Taken together, this is evidence that vowels in underlying onset clusters in Turkish lack durational and gestural targets. The present paper presents the ultrasound study, which provides direct evidence that vowels in Turkish onset clusters differ gesturally from phonologically targeted vowels, supporting the hypothesis that they result from an open transition between consonants.

2 Method

The present ultrasound study was inspired by Davidson & Stone (2003)'s use of ultrasound to determine whether articulation of /zg/ clusters by English speakers as [zəg] represents phonological epenthesis, or phonetic intrusion. They compared [zəg] to English words containing a lexical schwa (*succumb*), and words with sC clusters (*scum*), according to the logic that if the inserted vowel is epenthetic and has a gestural target, its gestural sequence will resemble lexical schwa's more closely than the insertionless cluster's. If the inserted vowel is instead intrusive, then its gestures will more closely resemble the insertionless cluster's.

Following the same logic, this study compares the articulation of underlying clusters /Cr/ in Turkish, with and without inserted vowels, to that of /CVr/ sequences, which contain underlying vowels.

2.1 Design

The primary independent variable in the experiment was the underlying structure of the target word, and hence the lexical status of the vowel between /C/ and /r/: non-lexical vowels occurred in words beginning with a stop+/r/ onset cluster

(/Cr/), and lexical vowels occurred in words beginning with a simple onset followed by an underlying vowel and /r/ (/CVr/). 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/[1]) were included. Due to a mistaken syllabification during the experimental design, the d-o condition did not contain a true onset cluster, and was therefore dropped from the analysis, leaving eight C1-V2 conditions; see Bellik (2019b: 39) for details. Both real and nonce words were included, to assess the productivity of insertion. No acoustic differences between insertion in real and nonce words were found, nor were any systematic gestural differences between inserted vowels in real and nonce words found, so I set aside the real/nonce distinction from here on.

Vowel insertion is reported to occur in /s/+stop and obstruent+/l/ clusters in Turkish, as well as obstruent+/r/ (Yavaş 1980, Clements & Sezer 1982). The Turkish /r/ is phonetically a voiced alveolar tap, while the /l/ can be a palatalized post-alveolar lateral or a velarized dental lateral, depending on context (Göksel & Kerslake 2005). This experiment uses stop+/r/ clusters because they have a higher rate of insertion in the Turkish Electronic Living Lexicon (TELL; Inkelas et al. 2000) than /sC/ clusters (71% vs. 42%). Also, 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; see also Bradley 2004). /Cl/ clusters were avoided for ease of segmentation using spectrograms.

Thus, the study had a 2 by 3 by 3 design (1).

(1) Experimental factors:
Word shape = [/Cr/ vs. /CVr/]
× [C1 = /b d g/]
× [V2 = /i a o/]

The full experiment, detailed in Bellik (2019b: ch. 2), included both careful and casual speech. The ultrasound results reported here are restricted to careful, hyper-articulated speech; see Bellik (2019a) for acoustic differences across speech styles.

[1]Originally /u/ was included as the third V2 value, rather than /o/, but no familiar words of the shape /bru-/ could be found, and so /o/ was selected.

2.2 Materials

A list of real and nonce words beginning with stop+/r/ clusters was generated. All onset clusters in Turkish occur in borrowed words; a familiarity-rating task with three Turkish speakers ensured that all real words in the study were familiar, not novel. Control words of the form /CVrV/ were created for every condition so that non-lexical vowels in /Cr/ words could be compared to lexical vowels in the same context. The apparent insertion of [ɯ] is attested before all qualities of V2, so /CɯrV/ controls were included for every V2 condition. In addition, /Ciri/ and /Curo/ controls were included, since insertion of [i] and [u] is reported before /i/ and /o/, respectively. Most control words are nonce words,[2] although real words were included where possible. Table 3 shows the experimental items and controls. See Bellik (2019b: Ch.2) for further details, including a list of the seventeen fillers.

Participants were instructed to speak carefully. All words were presented in a carrier sentence (2), which includes slots for two target words (X and Y). The sentence was designed to elicit contrastive focus on the target words, to promote hyperarticulation.

(2) Carrier sentence:
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.'

2.3 Participants

Seven native speakers of Turkish (4 female: S1, S4, S5, S7; age range 18–35) were recruited from the University of California at Santa Cruz. S1 participated in the pilot experiment with voiceless rather than voiced obstruents, and S2's ultrasound data were uninterpretable for anatomical reasons, leaving five speakers' ultrasound data to be discussed here. S3 is bilingual in French and Turkish. 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.

[2] Surface productions of /#CrV/ sequences often sound like [CɯrV], both in this experiment and according to TELL. However, across all values of V2, underlying /#CɯrV2/ sequences are rare (Bellik 2019b: ch. 5). In fact, such sequences where V≠a are completely unattested. This suggests that Turkish phonology prohibits these disharmonic sequences of segments, and can be taken as distributional evidence that there is no vowel segment/gesture between /C/ and /r/ in underlying clusters.

Table 3: Stimuli for the production experiment

C1	V2	Experimental: /Cr/ Real, nonce	Controls: /CVr/
b	/**i**/	bri.fing ‘briefing’ bri.mi.ti	bɯ.ri.pis bi.ri.m-in ‘unit.your’ bi.ri.bis
	/**a**/	bran.ʃ-ɯ ‘subject.ACC’ brat.ʧi.ten	bɯ.ran.ʤɯ
	/**o**/	bro.ʃyr ‘brochure’ bro.ʒør.le	bɯ.ro.ʒyn bu.ro.ʧyp
d	/**i**/	drip.ling ‘dribbling’ drip.li.ke	dɯ.rib.le di.rim.-ler ‘life.PL’ di.rib.rit
	/**a**/	dra.ma ‘drama’ dra.fa	dɯ.rap
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 (2.7) ‘gross.ACC’ gro.dol	gɯ.ron gu.rot

2.4 Procedure

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, 45–60 frames per second). Stimuli were presented to subjects on a laptop screen, one sentence at a time. Participants read through a list of 27 sentences five times in careful speech.

Acoustic recordings were annotated in Praat (Boersma & Weenink 2015) in order to identify the time range of the ICI. The left edge of the interconsonantal interval (ICI) was marked from the beginning of the C1 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/. Then the ultrasound frame best corresponding to the midpoint of the ICI was selected using a Python script, and tongue tracings were made in Edgetrak (Li et al. 2005).

R (R Core Team 2017) and the ssanova function in the gss package (Gu 2014) were used to create smoothing spline ANOVAs (SSANOVA; Gu 2002, Davidson 2005) for each word, within subject. An SSANOVA is essentially the mean of multiple curves, plus a confidence interval around it. Within each C1-V2 combination, the SSANOVAs for words with underlying clusters and with underlying vowels were plotted together.

2.5 Predictions

Lexical vowels have phonologically specified gestural targets, while intrusive vowels are targetless. Hence, the position of the articulators during an intrusive vowel will be determined by the gestural demands of the surrounding vowels and consonants. According to the hypothesis that the inserted vowels result from an open transition between consonants, the surrounding consonants and the following vowel should shape inserted vowels more than underlying vowels.

The phonological status of the inserted vowel shapes the organization of the syllable, which in turn shapes tongue position during the ICI. Let's take /gram/ [gɯram] 'gram' as an example. If the inserted vowel [ɯ] is epenthetic and has a gestural target, the word is syllabified as [gɯ.ram]. Thus, [gɯ] forms a syllable with a simplex onset and the gestures for [g] and [ɯ] are therefore coordinated with each other. The [r] and [a] gestures are part of a second syllable and are not directly coordinated with [g] and [ɯ], since they belong to distinct syllables, just as in a two-syllable /CVrV/ control word like /gɯrap/ [gɯ.rap] (nonce). This syllabification limits the overlap of the second vowel /a/ with the first vowel [ɯ]. On the other hand, if [ɯ] in [gɯram] is an intrusive vowel resulting from gestural timing alone, then /gram/ is a single syllable with a complex onset, meaning that the gesture for the nucleus /a/ is coordinated with both [g] and [r] gestures and is expected to overlap the interval between them (Byrd 1996, Browman & Goldstein 1988), pulling the tongue body toward /a/'s target during the ICI when [ɯ] is heard. We predict, then, that if a word-initial consonant cluster /Cr/ is syllabified as a complex onset with vowel intrusion, then the vowel that follows the consonant cluster will affect tongue position in the ICI in /Cr/ words more than in control /CVr/ words. But if a word-initial consonant cluster is broken up by epenthesis and syllabified as two consecutive syllables with simplex onsets, then this predicts that the vowel following the consonant cluster will equally affect experimental /CrV/ and control /CVrV/ words.

The degree to which the vowel following /Cr/ can overlap the ICI will be shaped by C's demands on the tongue body. Vowel-based differences between intrusive and underlying vowels are predicted to be most pronounced when the preceding consonant is labial, since the lips are able to move independently of the tongue body, meaning that an overlap of V2 with /b/ minimally interferes with /b/'s articulation. When the preceding consonant is coronal, its tongue tip target will limit movement of the tongue body toward V2's target, since the tongue tip is coupled to the tongue body, limiting V2's impact during the ICI. Finally, V2-driven differences between underlying and inserted vowels will be least pronounced when the preceding consonant is dorsal (/g/), since /g/'s tongue body target will most severely limit anticipatory movement toward the following vowel's tongue body target.

3 Results

The vowel that drives anticipatory coarticulation in the ICI determines the direction of expected differences between intrusive and underlying vowels. In the analysis that follows, I compare SSANOVAs of tongue body positions at the midpoint of the ICI for intrusive vowels with those for /ɯ/ (the most schwa-like Turkish vowel) and for the high vowels that regressive vowel harmony would demand, based on claims of harmony in the previous literature (Yavaş 1980, Clements & Sezer 1982, Kaun 2000, Yıldız 2010), within each /C+V2/ condition. Each speaker is plotted separately since there can be significant interspeaker variation. Only a representative sample of the SSANOVAs is included here; see Bellik (2019b: ch.2) for additional plots. Each plot shows a contour for tongue body position in the underlying cluster (orange line), another for underlying /ɯ/ (light blue), and a third for harmonic /i/ (dark blue). The legend shows the specific C(V)rV conditions for the plot, and the number of tokens that each curve represents. In most cases, there are half as many tokens of the condition where the V1 /ɯ/ is disharmonic with V2, because there are no real Turkish words with this shape; only nonce words were available. Where they have more tokens, the /Cr/ and harmonic V1 conditions combine the repetitions of a real and a nonce word.

The plots also include 99% confidence intervals, shown by dashed lines. When the confidence intervals for two curves do not overlap, the curves represent significantly different tongue positions. In most plots, the confidence intervals are so close to the main curve that they are hard to see. As is typical in ultrasound results, however, the position of the tongue tip and root is less certain, due to jaw shadow which limits imaging of those areas. This decreased certainty is reflected by the flared confidence intervals around the ends of the curves. The flaring at endpoints also reflects the method of calculating confidence intervals, which is sensitive to the square of the distance from the midpoint on the x-axis.

Because each SSANOVA curve represents only one to two lexical items from the same speaker, in interpreting the results, I did not consider the existence of a region of non-overlapping confidence intervals for a given pair of curves to be a sufficient basis for concluding that the articulation of the relevant cluster and /CVC/ sequence differed systematically across lexical items. I consider two curves to indicate a meaningful difference across conditions (not just lexical items) for a speaker only if there are statistically significant differences in tongue position across the majority of the length of the two curves and those differences are in the direction predicted by coarticulation with the surrounding consonants and vowel.

The results, detailed below, largely bore out the predictions above. Tongue position in underlying clusters differs systematically from tongue position in underlying vowels in the same context, in ways that show the greater influence of the following vowel on underlying clusters than on underlying vowels. Not every speaker conforms to the predictions in every condition, however. Also, a preceding /g/ obscures the effect of the following vowel on tongue body position.

Speakers fall into three groups, according to their acoustic and gestural results. The early bilinguals S3 and S6 tend to show the greatest differences between underlying clusters and underlying vowels. Late bilinguals S5 and S7 tend to show intermediate levels of gestural difference. Lastly, S4, the only monolingual in the study, has no acoustic differences between underlying clusters and underlying vowels (see Bellik 2018), but nonetheless displays gestural differences in most conditions.

3.1 /i/ conditions

When the following vowel is /i/, coarticulation will raise and front the tongue body in anticipation of /i/'s [+high, −back] target. Consequently, before /i/, the tongue body should be fronter and potentially higher during intrusive <v> than during underlying [+high, +back] /ɯ/. (Although both /i/ and /ɯ/ are phonologically [+high], Kılıç & Öğüt (2004) found that /ɯ/ has a lower tongue body position than /i/.) However, the tongue should be less front and high during a targetless <v> than during an underlying /i/: in an underlying /Cri/ sequence, /i/'s [+high, −back] target does not need to be attained until after both consonants, in contrast to an underlying /Ciri/ sequence. Therefore, if the vowel within the cluster is intrusive, we expect tongue body position in underlying clusters before /i/ to be intermediate between that of underlying /ɯ/ and /i/.

When the preceding consonant is /b/, four out of five speakers clearly bore out this prediction, as exemplified by S5 (Figure 2). The highest point in the tongue body at the midpoint of the ICI for /bri/ is intermediate in backness between that for /biri/ and for /bɯri/. For the last speaker (S4, Figure 3), the highest point of the tongue body in /bri/ is indeed much backer than in /biri/, but it is not fronter than the highest point in /bɯri/, which is markedly lower than the other two curves'.

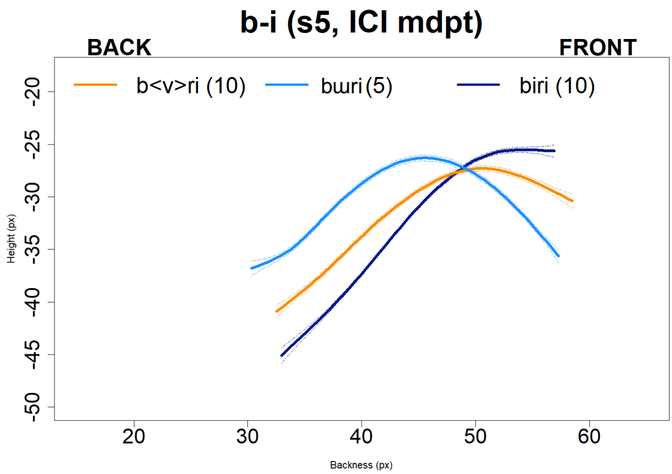


Figure 2: Tongue body position in the /b-i/ condition, S5

When the preceding consonant is /d/, the difference between tongue position in /i/ and /ɯ/ is less dramatic, likely due to the fronting effect of the coronal consonant (Figure 4). For all subjects, the peak of the tongue body during /dri/ is lower than its peak in one (S4, S7) or both (S3, S5, S6) of the underlying vowels – an effect not predicted by within-word coarticulation, although conceivably related to the final /a/ of the preceding word in the carrier phrase. For four out of five subjects, the confidence intervals around the /dri/ curve only overlap those for /dɯri/ or /diri/ when the curves intersect, and the peak of the /dri/ curve at the midpoint of the ICI is fronter than that of /dɯri/ but backer than that of /diri/. For three subjects, the tongue body's highest point during the ICI for /dri/ is intermediate in frontness between that of /diri/ and /dɯri/. For S7, however, the curve for /dri/ largely overlaps that of /dɯri/.

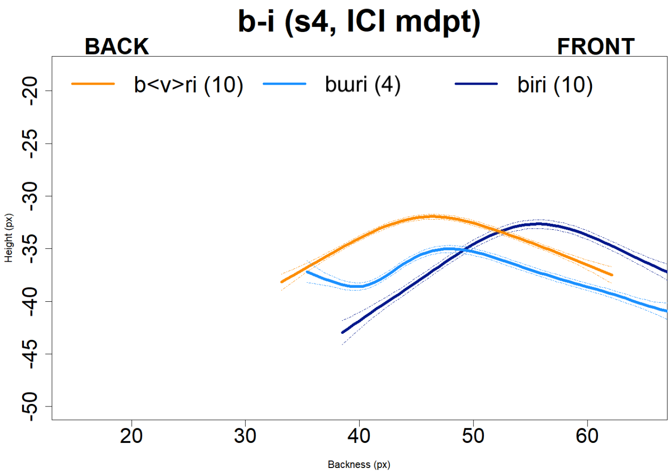


Figure 3: Tongue body position in the /b-i/ condition, S4

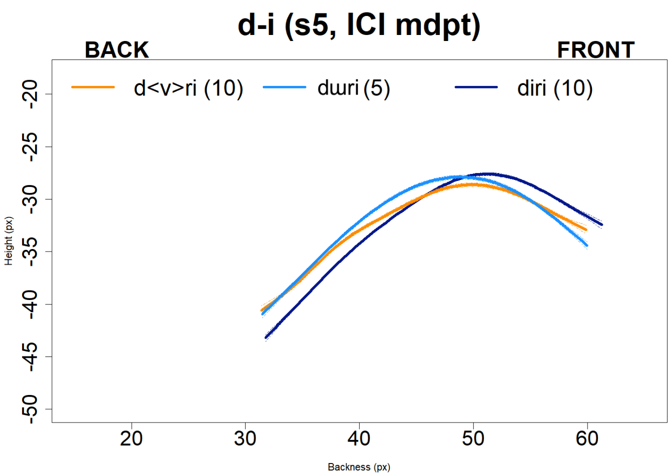


Figure 4: Tongue body position in the /d-i/ condition, S5

As expected, the pattern of clusters being intermediate between /i/ and /ɯ/ before /i/ is much less discernible when the preceding consonant is /g/ (Figure 5). The highest point of the tongue body during the ICI in /gri/ is intermediate in backness and height between that of /giri/ and /gɯri/ for S6 and S3, for both of whom /gri/'s peak is higher than that of /gɯri/, perhaps indicating that /g/ is having a greater impact during the targetless interval than during underlying

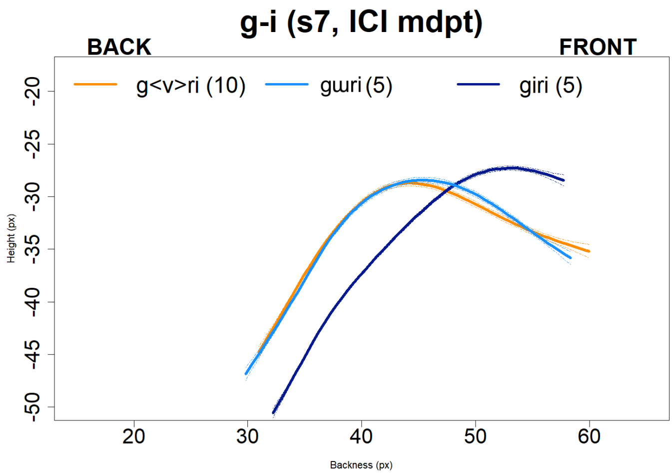


Figure 5: Tongue body position in g-i condition, S7

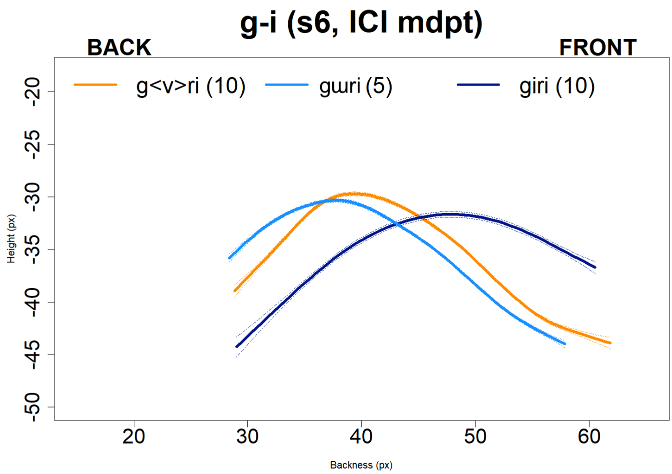


Figure 6: Tongue body position in g-i conditions, S6

vowels. 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). For S5, S7, and S4, meanwhile, tongue positions in /gri/ and /gɯri/ overlap for much of the curves' length.

To summarize results for the conditions where V2=/i/, tongue body position during the ICI in /bri/ and /dri/ is intermediate between that of /ɯ/ and /i/, while tongue body position in /gri/ is less clearly so, as expected.

3.2 /a/ conditions

When the following vowel is /a/, anticipatory coarticulation lowers the tongue body toward /a/'s [−high] target during the ICI. This predicts that a targetless vowel before /a/ will be lower than the [+high] /ɯ/. Since /ɯ/ matches /a/ in both backness and rounding, it is also harmonic with /a/, so no other vowels are predicted before /a/.

The prediction is most clearly borne out after /b/. Tongue position in the ICI for /bra/ is significantly lower than in /bɯra/ for all subjects at least at the peak and usually throughout the length of the curve (Figure 7). The height difference is greatest for early bilinguals S3 and S6.

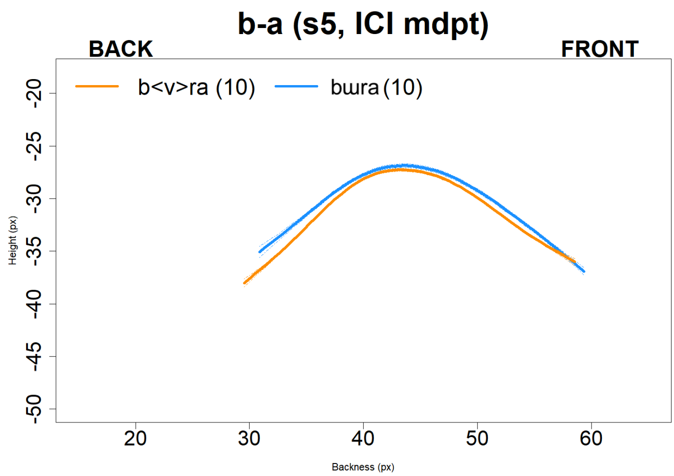


Figure 7: Tongue body position in the /b-a/ condition

The same pattern holds after /d/. The entire length of the tongue body, apart from any intersections, is significantly lower during the ICI in /dra/ than in /dɯra/ (Figure 8) for all subjects but S7.

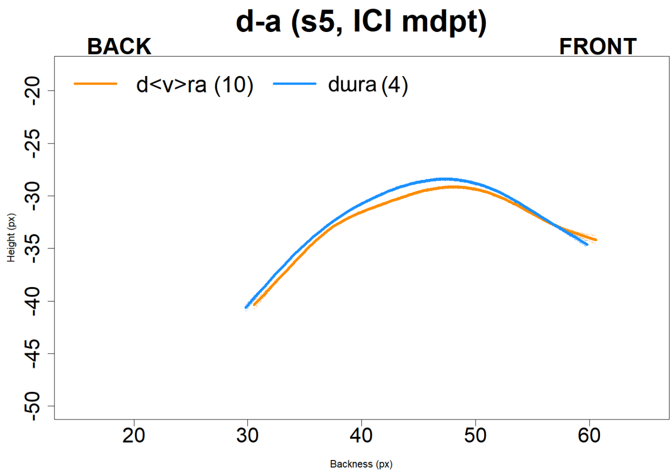


Figure 8: Tongue body position in the /d-a/ condition

S4 is the only subject who shows the predicted lowering in /gra/ (Figure 9). For S3, S6 and S7, the peak of the tongue body during the ICI in /gra/ is significantly higher than in /gura/, suggesting that /g/'s velar closure has a greater effect on inserted vowels for these speakers. For S4 and S6, the peak of the tongue body is fronter in /gra/'s ICI than in /gura/'s, perhaps reflecting movement toward /a/'s more central articulation.

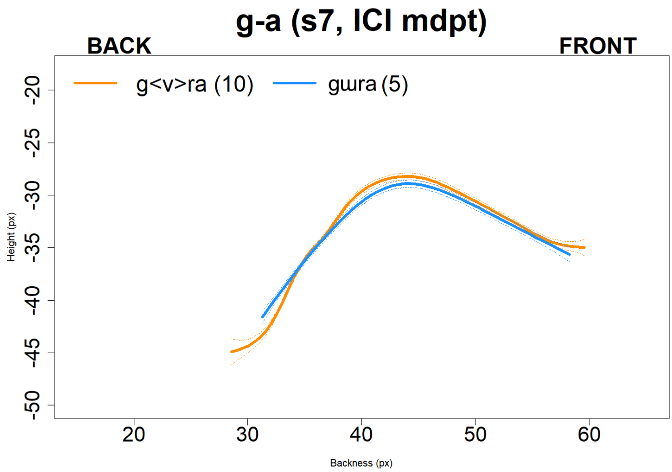


Figure 9: Tongue body position in the /g-a/ condition, S7 (S3 is similar)

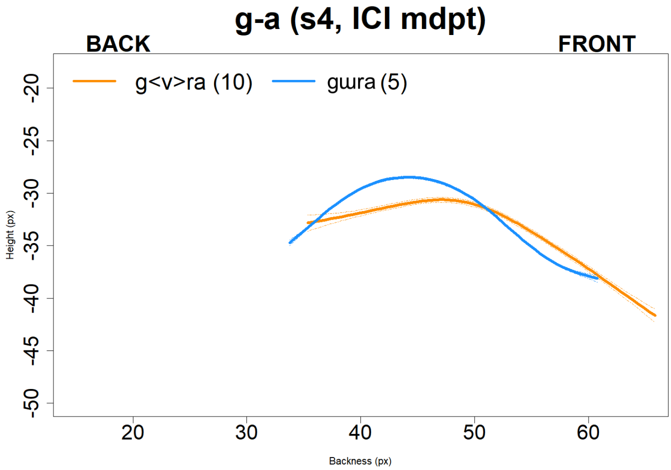


Figure 10: Tongue body position in the /g-a/ condition, S4 (S6 is similar)

Across consonant conditions, tongue position in clusters with a following /a/ tends to be lower than in an /ɯ/; however, a preceding /g/ usually blocks this effect. As in /i/ conditions, the following vowel /a/ has the greatest effect when /b/ precedes.

3.3 /o/ conditions

Anticipatory coarticulation with a following /o/ will move the tongue body toward a [+back, −high] target, possibly resulting in a lower tongue position for intrusive vowels than for high vowels /ɯ/ and /u/. No differences in backness during the ICI in /Cro/ vs. /Cɯro/ vs. /Curo/ are predicted. Differences in rounding are possible, but not observable from ultrasound.

Tongue body position during the ICI in /bro/ does tend to be lower than in /bɯro/ or /buro/ (Figure 11). For S3 and S7, the tongue body during the ICI of /bro/ is significantly lower than during both /bɯro/ and /buro/, at its peak and along most of its length. For S5 and S6, the tongue body position in /bro/ is significantly lower than in /bɯro/ but higher than in /buro/, while for S4, the curve for /bro/ is lower than /u/ but higher than /ɯ/ for most of its length.

A complicating factor is that S4 and S6 often fronted the back vowel preceding /o/, probably because the target words contained front rounded vowels in the syllable following /o/, in order to match the real word *broşür* /bro.ʃyr/ ‘brochure’. The nonce words in the condition were *brojörle* /bro.ʒør.le/, *buroçüp* /bu.ro.ʧyp/, *bırojün* /bɯ.ro.ʒyn/.

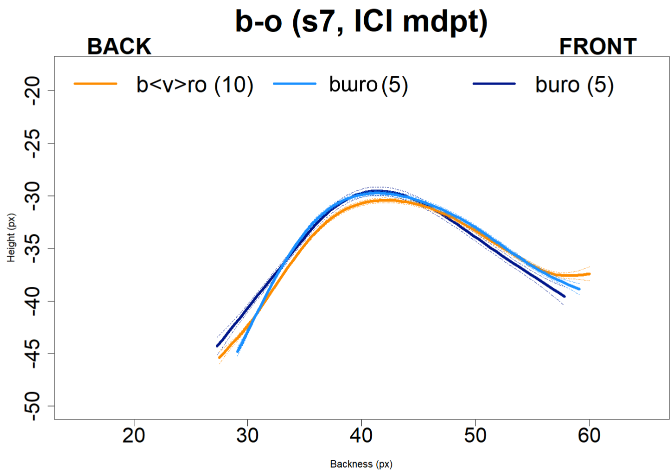


Figure 11: Tongue body position in the /b-o/ condition, S7

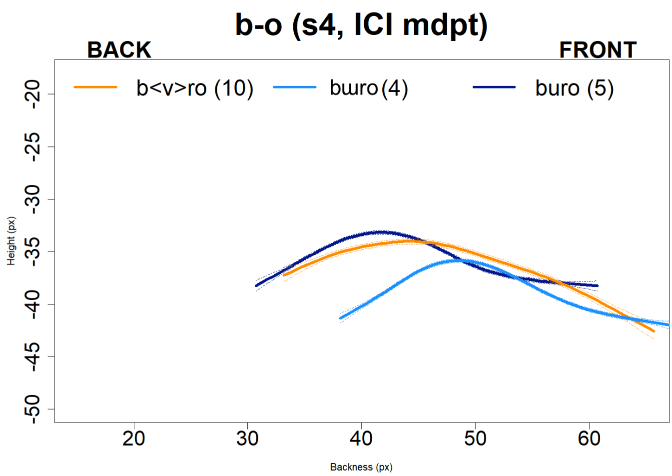


Figure 12: Tongue body position in the /b-o/ condition, S4

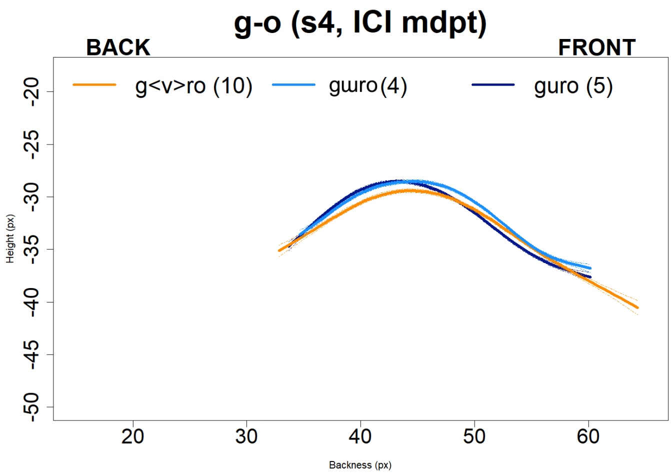


Figure 13: Tongue body position in the /g-o/ condition, S4

In the /g-o/ condition, most of the length of the tongue body during the ICI in /gro/ is significantly lower than in /gɯro/ and /guro/ for S4 (Figure 13), and lower than in /guro/ only for S7. For the other subjects, however, there are no significant differences in tongue body position between <v>, /ɯ/, and /u/ (Figure 14).

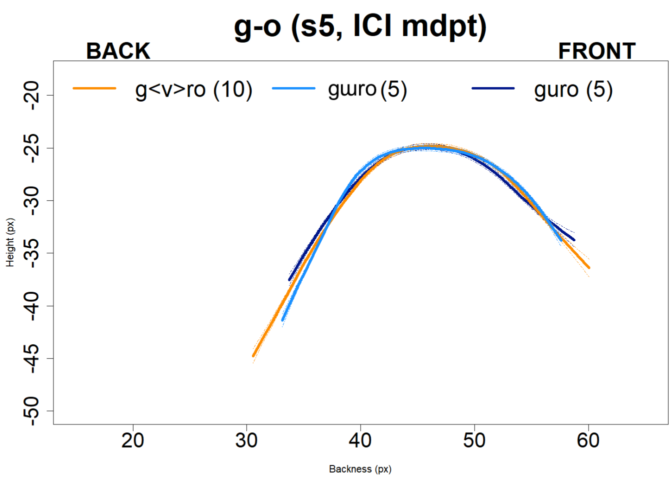


Figure 14: Tongue body position in the /g-o/ condition, S5

4 Discussion

This ultrasound study found that tongue body position during the ICI in underlying clusters differs significantly from that of underlying vowels. Generally speaking, anticipatory coarticulation with the following vowel appeared to influence tongue position in the ICI in /Cr/ more than in /CVr/. V2's influence was clearest when the preceding consonant was labial, whereas a preceding dorsal consonant largely blocks the lowering or fronting effect of the following vowel. This is expected, given that /g/ imposes its own target on the tongue body, unlike /b/ and /d/. These ultrasound results support the hypothesis that coarticulatory effects, rather than a gestural target, determine tongue position during acoustic vowels in complex onsets.

The tongue body positions found here indicate that, in underlying clusters, the tongue body is already moving toward its V2 target during the ICI. However, it has not yet attained the following V2's backness target, as shown by the fact that in /Cri/ words, the tongue is significantly backer during the ICI than it is during /Ciri/. Furthermore, the tongue has not attained its height target yet, since the tongue body is fairly high during the ICI even before a following mid or low vowel (/o/ or /a/). While a following non-high vowel does have a lowering effect, this effect is only clearly seen in the /b/ and to a lesser extent /d/ conditions. In the /g/ conditions, tongue body position during the ICI of clusters is not significantly lower than tongue body position in underlying high vowels for most subjects. This is particularly true for the /gro/ condition.

These gestural findings accord with the acoustic results in Bellik (2018) and Bellik (2019a), which found that the formant values of the ICI in /Cr/ words differed significantly from those of underlying vowels in the same context, particularly before /i/. Relatedly, the speech style comparison (Bellik 2019a) found that intrusive vowels before /i/ become more /i/-like in hypoarticulated speech, where greater gestural overlap is expected. A natural extension of the present study could examine the ultrasound data that were also collected in the hypoarticulated speech style condition to identify the gestural correlates of these acoustic results.

4.1 Gestural organization of Turkish onsets

These gradient distinctions between underlying vowels and underlying clusters, which vary according to the consonant context, imply that the gesture for V2 is still in its onset phase during the ICI; it has not attained its target yet. In the language of Gafos (2002) or Hall (2003), this suggests a gestural alignment in which the release of the C1 gesture is aligned with the onset of the V2 gesture. Indeed, if C1C2 is syllabified as a complex onset, then we would expect the V2 gesture to be coordinated with the C-center (Shaw et al. 2009, Browman & Goldstein 1988). Firmer conclusions about the relative timing of the gestures involved here, however, require a study of tongue movement over the course of the C(V)rV sequence, going beyond the analysis of tongue position at a single point in time (the midpoint of the ICI, or C-center, examined here). Optical Flow Analysis (Barbosa & Vatikiotis-Bateson 2014, Hall et al. 2015), for example, could indicate a stable point corresponding to each gestural target in the sequence, and show when the fastest vertical or horizontal changes in tongue position are occurring.

The gestural coordination that produces intrusive vowels seems to be grammaticized in some languages (Gafos 2002, Hall 2003), and this is likely the case in Turkish as well. One possibility is that the Turkish grammar of gestural timing prioritizes an anti-phase coordination between the two consonants in the cluster (which pushes them apart in time), over an in-phase coordination between C1 and V (which seeks to synchronize their onsets). The interspeaker variation found in this study suggests that Turkish speakers vary in the gestural coordination they employ in onset clusters. This variation is manifested in the acoustic characteristics of the ICI as well (Bellik 2018). This could reflect individual differences in the coupling strengths assigned to the competing gestures in the onset, as well as individual differences in phonetic implementation, although a study of a greater number of speakers is needed for any solid conclusions about the nature of interspeaker variation on this question.

4.2 Harmony and syllable structure

If onset cluster repairing vowels arise from gestural timing relations, then their behavior is irrelevant to 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. Its apparent harmonization (actually coarticulation) does not indicate that Turkish phonological vowel harmony can proceed from right to left, but its failure to harmonize also does not indicate that Turkish phonological vowel harmony is strictly limited to spreading from left to right.

Moreover, since the onset-repairing vowel is not epenthetic, it would seem that Turkish phonology does not categorically prohibit complex onsets in borrowed words, at least in bilinguals like the participants in this study. More work remains to be done, however, to uncover the role that experience with the source languages for these loanwords plays in Turkish speakers' grammatical restrictions on syllable structure. There is some initial evidence that these vowels do not have the same metrical status as full vowels, for a broader range of speakers, based on their text-setting (Bellik 2019b: ch.6).

4.3 Methodological contribution

Finally, this project bears on the extensibility of Davidson & Stone (2003)'s methodology, in adapting their experimental design but combining it with a more modern statistical technique, SSANOVAs. This comparative ultrasound methodology corroborated the acoustic study that probed the phonological status of vowels in Turkish onset clusters. This study finds a great deal of interspeaker variability in the articulation of these sequences. Further research with more speakers could illuminate factors that may structure this non-contrastive variability.

References

Barbosa, Adriano Vilela & Eric Vatikiotis-Bateson. 2014. *Optical flow analysis for measuring tongue-motion*. Paper presented at the 168th Meeting of the Acoustical Society of America, Indianapolis, IN.

Bellik, Jennifer. 2018. An acoustic study of vowel intrusion in Turkish onset clusters. *Laboratory Phonology* 9(1). 1–16.

Bellik, Jennifer. 2019a. The effect of speech style and deaccentuation on vowel intrusion in Turkish complex onsets. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 3498–3502. Melbourne.

Bellik, Jennifer. 2019b. *Vowel intrusion in Turkish onset clusters*. University of California. (Doctoral dissertation).

Boersma, Paul & David Weenink. 2015. *Praat: Doing phonetics by computer*.

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

Browman, Catherine P. & Louis Goldstein. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45. 140–155.

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

Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24. 209–244.

Clements, George N. & Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In H. van der Hulst & N. Smith (eds.), *The structure of phonological representations*, 213–255. Dordrecht: Foris.

Clements, Nick. 1991. Place of articulation in consonants and vowels: A unified theory. *Working Papers of the Cornell Phonetics Laboratory* 5. 77–123.

Davidson, Lisa. 2005. Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics* 19(6-7). 619–633.

Davidson, Lisa & Maureen Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. *Proceedings of the West Coast Conference on Formal Linguistics* 22. 165–178.

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

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

Gu, Chong. 2002. *Smoothing spline ANOVA models. Springer series in statistics.* New York: Springer-Verlag.

Gu, Chong. 2014. Smoothing spline ANOVA Models: R package gss. *Journal of Statistical Software* 58(5). 1–25. https://www.jstatsoft.org/v58/i05/.

Gudschinsky, Sarah C., Harold Popovich & Frances Popovich. 1970. Native reaction and phonetic similarity in Maxakali phonology. *Language* 46. 77–88.

Hall, Kathleen Currie, Claire Allen, Kevin McMullin, Veronica Letawsky & Alannah Turner. 2015. *Measuring magnitude of tongue movement for vowel height and backness.* Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015.

Hall, Nancy. 2003. *Gestures and segments: Vowel intrusion as overlap*. University of Massachusetts Amherst. (Doctoral dissertation).

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Inkelas, Sharon, Aylin Küntay, Ronald Sprouse & Orhan Orghun. 2000. Turkish electronic living lexicon (TELL). *Turkic Languages]* 4. 253–275.

Kabak, Bariş. 2011. Turkish vowel harmony. In M. Oostendorp, C. J. Ewen, E. Hume & K. Rice (eds.), *The Blackwell companion to phonology.* Hoboken: Wiley. DOI: 10.1002/9781444335262.wbctp0118.

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

Kılıç, Mehmet Akif & Fatih Öğüt. 2004. A high unrounded vowel in Turkish: Is it a central or back vowel? *Speech Communication* 45. 143–154.

Li, Min, Chandra Kambhamettu & Maureen Stone. 2005. Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics* 19(6-7). 545–554.

Padgett, Jaye. 2011. Consonant-vowel place feature interactions. In Marc van Oostendorp, Colin. J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 3, 1–26. Hoboken: Wiley. DOI: 10.1002/9781444335262.wbctp0075.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna. https://www.R-project.org/.

Shaw, Jason, Adamantios Gafos, Philip Hoole & Chakir Zeroual. 2009. Syllabification in Morrocan Arabic: Evidence from patterns of temporal stability in articulation. *Phonology* 26. 187–215. DOI: 10.1017/S0952675709001754.

Wrench, Alan. 2008. *Articulate assistant user guide*. version 1.17.

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

Yıldız, Yasemin. 2010. *Age effects in the acquisition of English onset clusters by Turkish learners: An optimality-theoretic approach.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Chapter 8

Intrusive and epenthetic vowels revisited

Nancy Hall
California State University Long Beach

The distinction drawn in Hall (2006) between epenthetic and intrusive vowels has been widely used in descriptive and theoretical work, and also challenged. This paper reviews some of the work that has engaged with the theory, in agreement or disagreement. This includes experimental approaches; proposed extensions to vowels inserted in other positions; and work arguing that vowels differ along a continuum rather than falling into these two classes. I argue that the dichotomy of presence/absence of a gesture is still a useful concept, while acknowledging that this is not the only important distinction among inserted vowels. An expanded typology recognizes subclasses among both intrusive and epenthetic vowels, as well as the possibility of alternation between epenthesis and intrusion.

1 Introduction

Vowel insertion is a cross-linguistically ubiquitous phenomenon – it is doubtful that any language does not display it in some way, at least in loanword adaptation or learner speech. But vowel insertion processes are diverse, in both their phonetic substance and their phonological patterning. Inserted vowels may be phonetically indistinguishable from lexical vowels, or similar but not identical to lexical vowels, or short, variable and indistinct. Some vowel insertion processes are optional, others obligatory. In terms of phonological behavior, some inserted vowels are clearly incorporated into the syllabic and metrical structure of the word, interacting with allophony, stress assignment, and intonation, while other inserted vowels are ignored by some or all aspects of the phonology. Some are fully accessible to the metalinguistic consciousness of speakers, as reflected

Nancy Hall. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542

for example in spelling, while in other cases native speakers seem unaware that they produce a vowel, despite it being clearly audible to linguists and visible on spectrograms. Describing and explaining this diversity is one of the tasks of phonological theory.

In Hall (2006), I argued that one important distinction among vowel insertion processes is whether the perceived vowel corresponds to a distinct articulatory gesture. This analysis is framed in the theory of Articulatory Phonology (Browman & Goldstein 1992b), in which abstract articulatory gestures are units of phonological representation. In VOWEL EPENTHESIS, speakers insert into the phonological representation a new vowel gesture, which was absent underlyingly and/or historically. In VOWEL INTRUSION, no new gesture is inserted, but a low degree of overlap between adjacent consonant gestures produces an open transition that can sound like a vowel. The class of intrusive vowels includes the short and indistinct vowels often called "excrescent" (Levin 1987), but also includes some vowels that are relatively long and distinct in quality.

This proposal was partly spurred by the popularity of typological studies of insertion in Optimality Theory (e.g., Kitto & De Lacy 1999, Fleischhacker 2005, among many others). Cross-linguistic variation along particular dimensions, such as inserted vowel quality or location of insertion, was often cited as evidence for universal constraints. However, some typologies mixed together insertion processes that, on reading of the descriptive sources, seemed like qualitatively different phenomena. In the push to capture all of phonology with a single set of constraints, the distinction between more and less "phoneticky" processes was sometimes ignored. The epenthesis/intrusion distinction was a way to bring this back.

Since then, the intrusive/epenthetic distinction has since been widely used in the analysis of vowel insertion phenomena. Examples of languages claimed[1] to have intrusive vowels include Albanian (Canalis 2007), Alsea (Buckley 2007), Qaqet (Tabain & Hellwig 2022), Khmer (Butler 2015), Norwegian (Garmann et al. 2021), ancient Carian (Adiego 2019), Pnar (Ring & Gruber 2014), as well as a large number of examples in Easterday (2017)'s study of highly complex consonant clusters. Examples of languages claimed to have both epenthetic and intrusive vowels (with different characteristics) include Tripolitanian Libyan Arabic (Heselwood et al. 2015), Turkish (Bellik 2019, 2024 [this volume]), SENĆOŦEN (Leonard 2007), and Sephardic Hebrew (Pariente 2010). Sometimes intrusion vs

[1]These are offered as examples of how the theory has been applied; I do not necessarily agree with all of the classifications mentioned in this paragraph. Also, note that some of these sources use the term "excrescent vowel" rather than "intrusive vowel", but adopt an explicitly gestural definition of excrescence that is in line with what I call intrusion.

epenthesis is posited as a difference between closely related dialects, as in Cavirani (in preparation)'s study of Carrarese and Pontremolese Italian. The distinction has also been applied to vowel insertion in L2 learner speech (Nogita & Fan 2012) and aphasic speech (Buchwald et al. 2006).

However, the dichotomy has also been challenged as too rigid, too simplistic, or simply not a useful framework. Blevins & Pawley (2010) argue that "a simple two-way division between intrusive phonologically invisible vowels and epenthetic phonologically visible vowels is too restrictive", preferring to classify inserted vowels primarily by their historical pathways of development. Hammond et al. (2014) suggest there is a continuum of vowel types, rather than a categorical division.

In this paper I review some of the recent work that has engaged with the theory – whether in agreement or disagreement. I argue that the dichotomy of presence/absence of a gesture is still a useful concept, and can explain some of the variation among inserted vowels, while acknowledging that this is not the only division in the typology.

The structure of the paper is as follows: Section 2 reviews the difference between vowel intrusion and epenthesis in Articulatory Phonology. Section 3 discusses recent experimental approaches to describing inserted vowels, and the challenges these methods may pose. Section 4 lays out an expanded typology that includes subclasses of intrusive and epenthetic vowels, as well as variation between epenthesis and intrusion. Section 5 discusses work that has extended the concept of vowel intrusion to vowels inserted word-initially, word-finally, or in VC transitions.

2 Intrusion vs epenthesis

The basic difference between vowel intrusion and vowel epenthesis is depicted in Figure 1. Both start with an underlying consonant cluster. In epenthesis, a new vowel gesture is added to the representation. The new vowel acts as a syllable nucleus, changing the syllabification of the consonants, with potential knock-on effects for other aspects of phonology that refer to syllables, such as stress assignment or minimal word requirements. Intrusion, on the other hand, occurs when the two consonants remain phonologically a cluster, but their gestures are "underlapped", producing an open transition that sounds vowel-like. This open transition may enhance the perception of both consonants.

The cluster with intrusion is nearly identical in representation to a "plain" cluster. Both have the same number of gestures, segments, and syllables; they differ only in the timing of the consonant constrictions. However, on an acoustic and perceptual level, the cluster with vowel intrusion is more similar to the cluster with vowel epenthesis. Both have a period of open vocal tract between the two consonant constrictions, which will sound vocalic. This mismatch between articulatory and perceptual characteristics is what can make vowel intrusion confusing: the vowels are "there and not there", as Ring & Gruber (2014) put it.

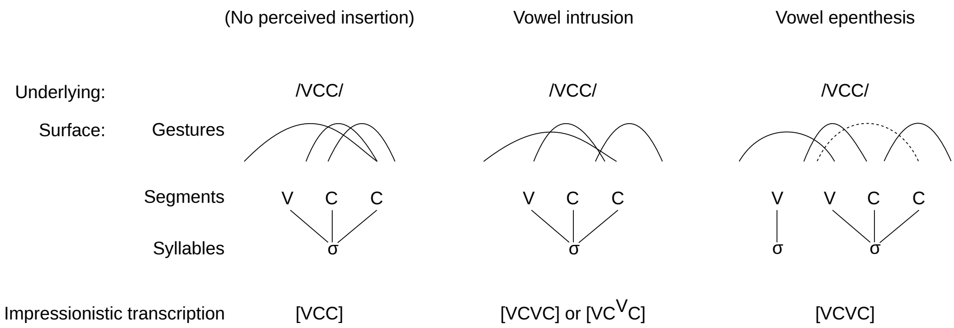


Figure 1: Gestural representations of consonant clusters without intrusion, consonant clusters with intrusive vowels, and epenthetic vowels. Dotted line = inserted gesture.

The different natures of these vocoids are reflected in different phonological patterning, different phonetic characteristics, and differences in speaker intuitions. Hall (2003, 2006) argues, based on a typological survey, that vowels which are plausibly intrusive tend to have a cluster of properties in common, as summarized in Table 1.

This clustering of properties can perhaps be best seen when both intrusive and epenthetic vowels co-occur in a single language. Two examples on which there has been recent work are Turkish and Tashlhiyt Berber.

Turkish has been described as inserting vowels into loanwords, both in CC onsets, as in [buluʒin] 'blue jeans', and CC codas, as in /sɑbr/ [sɑbɯr] 'patience'. Bellik (2019) argues that the vowels inserted in onsets are primarily intrusive, while those inserted in codas are epenthetic. The vowels in onsets are optional, and their acoustic quality is gradient and variable, with strong coarticulatory effects of nearby segments. Their durations are shorter than those of lexical vowels. They are not written in standard orthography. Bellik's ultrasound study finds that the tongue position is compatible with a "targetless" vowel. The vowels inserted in CC codas, on the other hand, seem to be true epenthetic vowels. They have a syllable repair function, occurring in marked, rising sonority coda clusters.

Table 1: Typical characteristics of intrusive and epenthetic vowels (adapted from Hall 2003, 2006)

Intrusive vowel	Epenthetic vowel
Phonetic and distributional properties	
Quality is transitional: either schwa, or influenced by surrounding consonants and vowels	Quality is grammatically determined; may be fixed or copied. A fixed-quality vowel does not have to be schwa
If the vowel copies the quality of another vowel over an intervening consonant, that consonant is a sonorant or guttural	If the vowel's quality is copied, there are not necessarily restrictions as to which consonants may be copied over
Likely to occur only in heterorganic clusters (or near taps)	Occurs in marked clusters, including homorganic ones
Likely to occur only if an adjacent consonant is voiced; voiceless clusters in the same language may have "aspiration"	Likely to have no specific association with voicing
Likely to be optional, have a highly variable duration and/or disappear at fast speech rates	Presence is not dependent on speech rate
Interaction with phonology	
May enhance perception in clusters, but does not have the function of repairing illicit clusters. May occur in relatively unmarked clusters	Repairs a marked structure, in the sense of something cross-linguistically rare, or avoided by other means in the same language
Does not form the nucleus of a syllable, cannot count for syllable-counting prosody	Forms the nucleus of a syllable
Meta-linguistic awareness	
Speakers may be unaware that they produce the vowel, and are unlikely to write it	Speakers will be aware of the vowel's presence

They are obligatory, written in the orthography, receive stress, and participate in a categorical phonological process of vowel harmony. While superficially, the insertion of vowels in onsets and codas might appear symmetrical in transcriptions, calling on the intrusion/epenthesis dichotomy helps to capture this cluster of ways in which the two phenomena differ.

A rather different example of a language that seems to show both vowel insertion types is Tashlhiyt Berber. Tashlhiyt allows long sequences of consonants, which are realized phonetically with intervening schwa-like vowels, as in /t-bdg/ [təbədəg] 'she is wet'. Since at least Dell & Elmedlaoui (1985), some linguists have argued that this schwa is a phonetic transition rather than a segment. As summarized by Ridouane & Cooper-Leavitt (2019: 434):

> Native speakers are largely unaware of the presence of this schwa in their speech, it does not affect intuitions about syllabification, it does not contribute to syllable weight in versification of traditional songs and it is ignored by phonological processes such as regressive obstruent devoicing.

It also shows distributional characteristics typical of an intrusive vowel, such as not appearing within homorganic clusters or voiceless clusters (where the transition is likely to be devoiced and hence non-vowel-like). In addition, Tashlhiyt has a longer schwa, whose insertion is prosodically triggered. Long schwa appears when segmental material is needed to allow expression of a complex final intonational melody in an emphatic statement, as in /imsʁ/ [imsəːːʁ] 'he erased'. Ridouane & Cooper-Leavitt (2019) argue that the two schwas should be seen as separate elements, which they term T-vocoid (transitional) and P-vocoid (prosodic). The T-vocoid appears to be intrusive, and the P-vocoid epenthetic, albeit driven by intonation rather than syllabification.

3 Experimental approaches

Hall (2003, 2006) drew largely on impressionistic written descriptions to identify likely cases of vowel intrusion. At that time, relatively few vowel insertion phenomena had been studied in the lab, with the notable exceptions of Scots Gaelic (e.g., Ladefoged et al. 1998, Bosch & De Jong 1997) and Dutch (e.g., Van Donselaar et al. 1999, Warner et al. 2001). Since then, there has been an increase in the use of experimental methods to test the phonetic and psychological nature of inserted vowels. New studies have greatly expanded the empirical basis for analyzing vowel typology. However, interpretation of experimental results is not always straightforward, and points to a need to continue developing theories of how abstract phonological representations relate to acoustic characteristics and speaker behaviors.

The gold standard of evidence for a gestural representation is direct imaging of articulation. A few studies have used ultrasound to test whether tongue position during an inserted vowel was consistent with a purely transitional trajectory (intrusion), or whether there was evidence of the vowel having its own target constriction (epenthesis). Davidson & Stone (2003) examined English speakers' nonce productions of pseudo-loanwords with illegal clusters, such as *zgomu* [zəgomu], and concluded that the schwas they inserted were transitional. Bellik (2019) drew the same conclusion regarding vowels inserted in real Turkish loanwords (see also Bellik 2024 [this volume]). On the other hand, Buchwald et al. (2006) found that schwas inserted in the speech of an aphasic did show evidence of having a distinct target. Such studies unfortunately remain rare, due to the expense and difficulty of articulatory imaging.

3.1 Acoustic measures

Acoustic studies of vowel insertion are far more common, and have taken a variety of approaches to testing whether a vowel is likely to be associated with a gesture. A common finding is that putative intrusive vowels have shorter durations than lexical schwa or other unstressed vowels (see Ring & Gruber 2014 for Pnar, Garmann et al. 2021 for Norwegian, Bellik 2019 for Turkish). It is not clear that this is always true, however. For example, Tabain & Hellwig (2022: 14) find that excrescent schwa in Qaqet has a duration comparable to that of phonemic schwa, 57 vs 56 ms, as well as similar formant values. (The evidence for its excrescent status comes primarily from speaker intuitions and phonological patterning.)

As noted in Section 2, intrusive vowels in voiced clusters are predicted to correspond to voiceless transitions (often called "aspiration") in voiceless clusters. Both result from the same gestural phasing. Accordingly, some studies examine voiced and voiceless transitions together. In Khmer, for example, Butler (2015) compares words containing vowels or aspiration that she suspects to be intrusive (C_CVC, where _ is the location of a voiced or voiceless transitional element), against CʌˈCVC disyllables and CʌC monosyllables. She finds that the voicing of the transition is predictable from the voicing of the preceding consonant, and that voiced and voiceless transitions have the same average duration, which is shorter than unstressed lexical [ʌ]. The formant values of the transitional vowels are also significantly different from those of [ʌ], and more influenced by the place of the preceding consonant. She concludes that the schwa-like sounds heard in /CCVC/ words are indeed transitions within monosyllables, and that consequently "the distinction between monosyllables and disyllables in Khmer is more clear-cut than previously thought." She also suggests that the typological concept of "sesquisyllables", as applied across languages to words with initial Cə-elements, has conflated structurally distinct word types.

In cases where it is difficult to directly compare inserted and lexical vowels in equivalent contexts, sometimes each can be compared to some point of reference in the same word. Karlin (2021) uses this technique in a small-scale corpus study of vowel insertion in Finnish archival recordings. She compares the vowels inserted in certain medial CC clusters, as in *silmä* [silimä] 'eye', against words with lexical vowels in the same position, such as *niminen* 'so-named'. To control for the variety of speakers and prosodic contexts in the corpus, the measures used were the duration ratio of V_2 (inserted or lexical) to the lexical V_1 of the same word, and the Euclidean distance between the qualities of V_2 and V_1. By both measures, vowels inserted in most /CC/ clusters patterned with lexical vowels in the same position, suggesting they are not synchronically intrusive (contra impressionistic descriptions of speaker intuitions from Harms 1976: 74 and Wiik 1965: 28). However, in one dialect the vowels in /hC/ clusters did show the short duration and variable presence that are characteristic of intrusion.

While the studies above compared inserted to lexical vowels, acoustic measures can also reveal diversity among inserted vowels. For Tripolitanian Libyan Arabic, Heselwood et al. (2015) examined inter-consonantal intervals (ICIs) in sequences of 2–4 stops, in phrases like /fak#tkasir/ 'jaw broke'. Previous descriptions had suggested that vowel epenthesis is possible in any CC sequence. The researchers find that on a range of measures, ICIs fall into two groups. Their durations have a bimodal distribution, with peaks around 20 ms and 50 ms. The shorter ICIs have the characteristics of intrusion: they tend to disappear at fast speech rates, and are often voiceless when between voiceless consonants. Short ICIs do not block voicing assimilation between flanking consonants. The longer ICIs have characteristics of epenthesis: they are more consistently present, more consistently voiced, and block consonant voicing assimilation. Short and long ICIs are generally associated with different insertion sites, although one context has both (as discussed further in section 4.3).

3.2 Challenges in interpreting acoustic results

The interpretation of acoustic results can be ambiguous, particularly when studies look beyond the inserted vowel itself to effects elsewhere in the word. Vowel insertion can be associated with durational adjustment throughout the word, not always in ways that are expected. For example, when Van Donselaar et al. (1999:

65) recorded the same Dutch words with and without schwa insertion to use as perception stimuli (e.g., *tulp* ~ *tul*[ə]*p* 'tulip'), they were surprised to find that the tokens with schwa insertion were shorter than those without. Hickey (1985) similarly claims that vowel insertion in Irish English codas cooccurs with shortening of the preceding vowel and sonorant, so that dialectal variants like [fɪːlm ~ fɪlɪm] *film* have the same overall duration. Gick & Wilson (2006) present similar findings for optional schwa insertion between high tense vowels and liquids in English: the rimes of words like *hee*(ə)*l* have the same duration with or without the schwa. They interpret this as evidence that the schwa is transitional.

However, it is not always clear which global durational adjustments are consistent with gestural retiming (intrusion) versus gesture insertion (epenthesis), nor exactly how each phenomenon would interact with other duration-affecting factors such as pre-boundary lengthening, polysyllabic shortening, or speech rate variation. Interpretations typically rely on a researcher's intuitive understanding of how different gestural structures would be realized, not backed by computational simulations within a spelled-out model of speech timing (like those of Gafos 2002 or Browman & Goldstein 1992a).

An example of case where interpretations might differ is Smith (2019)'s acoustic study of schwa insertion in Scottish English codas. Smith establishes that the inserted schwas in words like *form* [forəm] are shorter than the underlying vowels in words like *forum*, averaging 54 versus 95 ms respectively, and that the inserted schwas' duration positively correlates with that of other sounds in the word. Smith interprets both facts as evidence that the vowels are epenthetic rather than intrusive. She reasons that purely transitional vowels would be less than half the duration of underlying vowels, and that

> if the inserted vowel is an articulatory byproduct of low degree of overlap between the gestural phases of the surrounding consonants, its duration should not vary with duration of the surrounding segments, but remain uniformly short (Smith 2019: 133).

To me, neither assumption is obvious. The articulatory simulations in Browman & Goldstein (1992a: 55) show an example of a targetless (i.e., intrusive) schwa with duration well over 50 ms, and it seems conceivable that slower speech could be associated with lower gestural overlap, producing longer transitions. I also find it interesting that Smith reports greater rates of schwa insertion in CC codas (*farm*) than in CCC codas (*farms, farmed*). This seems opposite what one would expect for structure-repairing epenthesis, which ought to preferentially target the most marked syllables. On the intrusion analysis, on the other hand, it seems conceivable that CCC codas are produced with greater gestural crowding and overlap than CC codas, which would decrease the chance of open transition.

However, these differences in interpretation must remain a matter of opinion until a specific hypothesis about the gestural structure of Scottish complex codas is simulated and tested against natural data. There is a great need for more articulatory simulation work to clarify what phonetic patterns can and cannot be produced through retiming alone.

3.3 Speaker intuitions and metalinguistic tasks

Since intrusive vowels are not phonological units, they are expected to be less salient to speakers' conscious awareness than regular consonants and vowels. A number of studies have probed speaker intuitions concerning inserted vowels, and frequently find them to be treated differently than lexical vowels. Yet interpretation of these differences can be difficult, due to unclarity about what level of representation various tasks draw on.

The most basic question is whether speakers are even aware the vowel is there. This is often apparently assessed through discussion or questioning, although methods for eliciting these judgments are rarely described in detail. For example, Garmann et al. (2021) report of the very short vowels found in words like [bəɭɔː] 'blue' that "anecdotally, native speakers of Norwegian tend not to be aware of the intrusive vowel, and report difficulty in perceiving it even when the articulation is pointed out to them." Similarly, Karlin (2021) directly asked Finnish speakers whether vowel insertion is possible in various /rC/ clusters. She cites their lack of awareness as evidence for the intrusive status of these vowels (in /rC/ clusters specifically, unlike other CC clusters).

Speakers can also be directly asked for syllable count judgments, but responses may be contaminated by knowledge of standard orthography as well as prescriptive notions of the syllable. Bellik (2019: 164) recounts one conversation among Turkish speakers:

> Speaker M initially said that *prens* [pirens] and *kral* [kɯral] were disyllabic. Speakers S and E then asserted that M was wrong, and *prens* and *kral* were monosyllabic. S and E argued that the definition of a syllable is a vowel, and since *prens* and *kral* are spelled with only one vowel, they were by definition monosyllabic. Faced with this argument, M lost confidence in his judgement of bisyllabicity. The fourth speaker, N, was reluctant to take a side.

Another speaker claimed that *prens* has one and a half syllables. This phenomenon of speakers counting non-whole numbers of syllables (including in tasks where that wasn't intended to be an option) has also been reported for Scots Gaelic (Hammond et al. 2014), and for English words like *fire* (Cohn & Tilsen 2015). These stories suggest that speakers' definition of a "syllable" does not always align with the concept of a syllable in phonological theory. It is important to design instruments, like that of Cohn & Tilsen, that allow speakers to give intermediate responses, as that detail will be lost if they are forced into whole-syllable judgments. The question of what exactly speakers are counting when they talk about half syllables remains unclear. It could be a way of describing a non-syllabic phonetic sonority peak, such as an intrusive vowel.

Some studies, rather than asking directly, use tasks designed to indirectly access syllabicity judgments. Ring & Gruber (2014) used a language game (based on a method developed by Van Donselaar et al. 1999) to test the extra-short vowels that Pnar speakers optionally insert in onset clusters, such as [kba ~ kəba] 'rice'. Speakers heard audio recordings of real and pseudo words. They were told to reverse the phonemes in any word that was a monosyllable (tam → mat), but reverse the syllables of disyllables (batam -> tamba). Both real and pseudo words with inserted schwa were predominantly treated as monosyllables. Real words with lexical schwa in the first syllable were treated as disyllables, even when the schwa was clipped to the typical duration of an inserted vowel. This suggests that speakers have different representations of lexical and non-lexical schwa, and do not consider the inserted schwa syllabic, consistent with the predictions of the intrusion analysis.

In Turkish, Bellik (2019) examines whether intrusive and lexical vowels are treated alike in text-to-tune setting. In a corpus of recorded songs, she finds that vowels inserted in complex onsets can be set to a beat (48/82 tokens), but often are not (34/82), while lexical vowels in the same position are almost always set to at least one beat. There is variation across lexical items, where the words with the most consistent insertion are most likely to have a beat for the inserted vowel. However, an elicitation experiment had different results. When asked to insert words into a two-note slot, speakers either suppressed vowel insertion, or produced it and gave the intrusive vowel its own note. This is problematic for the intrusion analysis, given that the presence/absence of an inserted vowel changed the metrical setting (but see discussion in section 4.3). See also Baronian & Royer-Artuso (2024 [this volume]) for discussion of issues in interpreting text-setting as evidence in phonology.

Although standardized orthography can contaminate judgments, spontaneous or unstandardized orthography can be a useful indicator of speaker intuitions. For example, Qaqet has optional schwas, mostly adjacent to liquids as in [mraɽik ~ məraɽik] 'cross', as well as a phonemic schwa as in [aləm] 'feather' (Tabain &

Hellwig 2022). The vowels differ in phonological behavior, with only the phonemic schwa triggering lenition of voiceless plosives. Tabain & Hellwig (2022) claim that the optional schwas are excrescent rather than epenthetic, since they do not break illegal consonant clusters. Interestingly, they report that the two schwas are acoustically comparable in duration and formant values. However, speakers treat the two schwas differently:

> Speakers agree on the presence of the phonemic vowel when asked for an orthographic representation, but do not write the excrescent vowel, whose presence is indeed highly variable across speakers and across lexical items. (Tabain & Hellwig 2022: 14)

Other examples where speakers do not write plausibly transitional vowels include Lamkang (Burke et al. 2019) and Kalam (Blevins & Pawley 2010, discussed further in section 4.4).

A recurrent challenge in interpreting metalinguistic tasks is how to understand gradient results. For example, Hammond et al. (2014) used a battery of tasks to probe the intuitions of elderly Scots Gaelic speakers. Words with inserted vowels, such as *ainm* [ɛnɛm] 'name', were compared to words with only lexical vowels, such as *anam* [anam] 'soul'. Tests included identifying which of two words an audio (V)CVC sequence was clipped from, writing nonce words in Gaelic orthography, counting syllables, knocking syllables, and identifying the first and last syllable of the word. The syllable-identification task was intended to probe syllabification of the medial consonant, which has been reported to act as a coda before lexical vowels but an onset before inserted vowels. Speakers treated inserted vowels differently than lexical vowels on nearly every measure, but in every case showed intermediate and inconsistent results. For example, speakers counted the inserted vowel as a syllable 56% and 59% of the time in the counting and knocking tasks, as opposed to 82% and 87% for lexical vowels, and a number of participants tried to offer answers such as "one and a half syllables". Syllabification results were also puzzling: when asked to identify the first and second syllables of CVCVC words with inserted vowels, speakers tended to not to produce the medial consonant in either syllable, leaving its structural affiliation unclear. The authors conclude that the inserted vowels are phonological objects and not phonetic transitions, given that they affect syllabification and count as syllables at least some of the time for speaker intuitions.

In some cases, it can be debatable what level of linguistic structure a given speaker behavior taps into. For example, Burke et al. (2019) describe an ambiguous case in the Trans-Himalayan language Lamkang. Lamkang optionally inserts extra-short vowels into strings of consonants. Up to 4 consonants can occur word-initially, through prefixing, as in *mtknoolam* ‘they are shaking me off.’ Phonetically, such a word can be realized as [mᵊtᵊkᵊnoolɐm], with extra-short vowels between the concatenated consonants. Extra-short vowels average 33 ms in duration, compared to 85 ms for lexical short vowels. Although missionary translators wrote the vowels in the orthography, native speakers do not. These acoustic and orthographic characteristics seem consistent with the vowels being intrusive.

Yet when Burke et al. (2019) asked speakers to clap or tap to the syllables in such words, they clapped to each extra-short vowel. Variation in vowel insertion corresponded with variation in clapping. Among 6 speakers there were 4 pronunciations of the word /m-t-pmen/ ‘she is trapping me’, with 2–4 claps: [mᵊtpmɛn], [mᵊttᵊpmɛn], [mᵊtpᵊmɛn], and [mᵊttᵊppᵊmɛn]. Burke et al conclude from the clapping task that the extra-short vowel does act as a syllable nucleus, which would mean it is not intrusive.

However, the high variability in the responses seems atypical for syllable count judgments. This raises the question of whether the clapping task is actually tapping into intuitions about syllables in the phonologist’s sense, or something else. Clapping is known to reflect different structures in different languages; while English speakers clap to syllables, Japanese speakers clap to moras. It could be that clapping in Lamkang reflects surface intensity peaks, or that it taps into some type of timing structure that is used in coordinating long series of consonants.

In short, metalinguistic tasks are turning up fascinating data on speakers’ knowledge of lexical and inserted vowels, which require explanation. Yet interpretation of the facts can be difficult. There isn’t an explicit theory of how each metalinguistic task relates to phonological representations in each language. We don’t always know to what extent different tasks tap into knowledge of orthographic forms, underlying lexical forms, surface phonological representations, and/or purely acoustic forms. Methodological refinements may clarify some cases where a metalinguistic behavior seems to not match other indicators of transitional vs segmental status, such as Lamkang.

3.4 Orthography and historical development

Phonological representations can influence spelling, but Bellik (2019) has suggested that influence may also flow the other way: an orthographic convention of not writing a vowel could potentially reinforce speakers’ internal representation of the vowel as transitional, delaying the common diachronic path whereby intrusive vowels are eventually phonologized as segments.

Bellik (2019) points out that vowels in older Turkish loans seem to have been reanalyzed as segments and lexicalized fairly rapidly. Walter (2018) shows that in Ottoman-era texts, many loans from European languages were written down with an inserted vowel, and those vowels are lexicalized today. Yet Bellik (2019) argues that newer loans seem to be retaining vowel intrusion as a somewhat stable configuration. She suggests that the difference relates to speakers' metalinguistic knowledge. A younger generation of the Turkish speakers are aware of the prescriptive spelling of newer loans, with no orthographic vowel corresponding to the intrusive vowel. They are also likely to have studied English or other source languages for the loanwords, and hence be aware that the source words have a consonant cluster. Both factors increase the likelihood that they will continue to interpret the vocalic period in CC onsets as an open transition rather than a vowel segment. Bellik cites anecdotal evidence that children often include the vowel when learning to write, and are taught not to. This suggests that intrusive vowels in these loanwords would naturally tend to be on their way to reanalysis as segments, were it not for external factors.

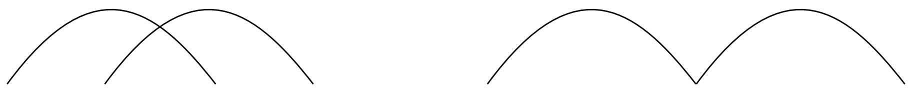
4 Expanding the typology

As discussed in Section 3, in-depth examination of any specific vowel insertion process by multiple measures tends to find less-than-perfect adherence to the list of prototypical qualities of intrusive and epenthetic vowels given in Table 1. Some linguists have concluded from this that, in the words of Hammond et al. (2014), "any phonology-phonetics distinction for inserted vowels is probably more of a continuum, rather than a categorical split, or at least that there are more than two types of inserted vowels." In this section I lay out one version of what a more articulated typology could look like. This proposal retains the basic dichotomy of presence/absence of a vowel gesture, but acknowledges three types of additional phonetic variation: 1) the existence of more than one gestural timing configuration that can produce vowel intrusion, 2) the existence of gesturally stronger and weaker types of epenthetic vowel, and 3) the possibility that what linguists describe as one vowel insertion process sometimes involves a mixture of epenthesis and intrusion. This typology offers more ways to categorize inserted vowels, while keeping the insight that a definite set of gestural facts need to underlie each type. The predictions of this theory remain more restrictive than appeal to a loosely-defined continuum.

Note that this typology does not attempt to address purely phonological diversity among true epenthetic vowels, such as variation in their quality, contexts, interaction with stress, etc. I assume that these characteristics are controlled by phonological constraints of some type, and are not limited by the vowels' phonetic substance, which is the focus here.

4.1 Subtypes of vowel intrusion

Intrusive vowels are not all the same, because they can result from different levels of consonant overlap. This was established by Gafos (2002), whose simulation work showed that some characteristics of the intrusive vowel are expected to change as the overlap decreases, as depicted in Figure 2 and (1). At a moderate degree of overlap (when the articulation of C_2 starts around the articulatory midpoint of C_1), transitions within consonant clusters will occur only at slower speech rates and in heterorganic clusters. Consonant voicing also has a strong influence at this level of overlap; the transition will be voiceless (often described as "aspiration") if the flanking consonants are voiceless. But all of these characteristics change if the consonants have little or no overlap, as shown on the right. In clusters of non-overlapping consonants, release will also occur even if they are homorganic, even at faster speech rates, and even in voiceless environments.



(a) Open transition in Cs with moderate overlap (b) Open transition in non-overlapped Cs

Figure 2: Subtypes of vowel intrusion

(1) Subtypes of vowel intrusion
  a. Open transition in Cs with moderate overlap
    - Disappears in fast speech
    - Occurs in heterorganic clusters only
    - Transition is voiced only near voiced C; realized as "aspiration" in voiceless environment
    - Example: Sierra Popoluca (Gafos 2002): [miɲᵊpaʔ]
  b. Open transition in non-overlapped Cs
    - Present in fast speech
    - May occur in homorganic clusters
    - May be voiced between voiceless Cs
    - Example: Moroccan Colloquial Arabic (Gafos 2002): [zˤnatˤᵊtˤ]

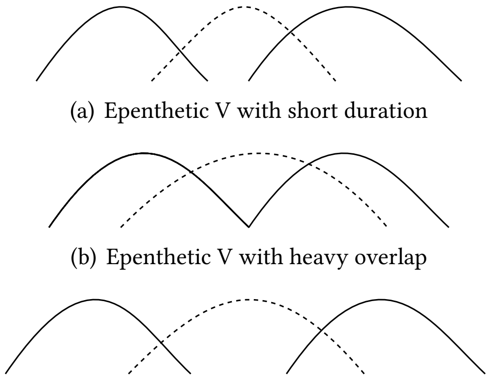
It is predicted that we should find some cases of vowel intrusion that occur in a wider variety of cluster types, and are less speech-rate dependent than others. This is one reason that cases of genuine intrusion could depart from the prototypical list of characteristics given in Table 1, and why certain characteristics in that table are hedged as "likely", rather than required. If a vowel has several characteristics of intrusion but not exactly the typical distribution pattern, it is worth considering that it could be the lower-overlap type.

4.2 Phonetic subtypes of epenthetic vowels

A second source of variation among inserted vowels is that when a new gesture is inserted, it is not necessarily always the same kind of gesture, or phased the same way with respect to other gestures. There is evidence for phonetic diversity among true epenthetic vowels. Some appear to be phonetically identical to lexical vowels in the same language, while others are not.

For example, Hall (2013) finds that some (but not all) Lebanese Arabic speakers produce epenthetic vowels that are shorter in duration and more centralized in quality than lexical vowels. Words like [libis] 'wore' (from /libis/) and [libis] 'clothes' (from /libs/) are not quite homophonous. Yet the inserted vowels are not merely transitional; they have all the characteristics of true epenthesis. They have a fixed quality, which varies regionally from [i] to [e] to [ə], and they can have a long duration in certain prosodic contexts such as phrase-final lengthening. Their presence is optional, but in an all-or-nothing way rather than a phonetically gradient way. Vowel insertion occurs most frequently in rising sonority CC codas, where it has a clear syllable-repair function. It can occur in homorganic clusters like /nt/, which are less prone to open transition. Speakers are consciously aware of the vowels, will write them if asked to write in the colloquial, and can suppress them when speaking in a formal register. All of this suggests that the inserted vowels have their own gestures, which are not necessarily identical to the gestures of similar lexical vowels.

I am not aware of any explicit proposal for how to model phonetically weak true epenthetic vowels, like those in Lebanese, in gestural representations. Weak epenthetic gestures could have special settings for properties such as stiffness, which affects duration, or blending strength, which affects how much two competing gestures influence an articulator's trajectory. Another possibility is that epenthetic vowels are more overlapped by nearby consonants than lexical vowels are. Some hypothetical gestural representations of weak epenthetic vowels are shown on the left of Figure 3.



(a) Epenthetic V with short duration

(b) Epenthetic V with heavy overlap

(c) Phonetically normal epenthetic vowel

Figure 3: Phonetic subtypes of epenthetic vowels. Hypothetical representations of phonetically weak and strong epenthetic vowels. (Dotted line = inserted vowel)

Epenthetic vowels are not always phonetically weak, however. Hall (2013) found that some Lebanese speakers have epenthetic vowels that are acoustically indistinguishable from lexical [i]. Other reports of phonetically normal epenthetic vowels include Mohawk (Michelson 1989: 40) and vowels in Korean loanwords (Kim & Kochetov 2011). This situation is depicted on the right of Figure 3.

Phonetic weakness in true epenthetic vowels could be a natural and fleeting stage of diachronic development, or it might persist for functional reasons. Epenthetic vowels carry no lexical information of their own. Keeping them phonetically reduced is in line with a general tendency to maximize informative material and minimize predictable material. It would be useful to have more studies of the phonetics of true epenthetic vowels, to better understand the range of variation within this category.

4.3 Variation between gesture presence and absence

A final complication for classifying vowels as intrusive or epenthetic is that there could be variation between gestural presence and absence. This possibility is depicted in Figure 4. Such variation could occur across phonological contexts, or across speakers, or across registers.

Alternation between an intrusive vowel and a segmental vowel is a logical possibility anywhere that segmental CC ~ CVC variation occurs. One piece of evidence that this actually does occur is Heselwood et al. (2015)'s acoustic study

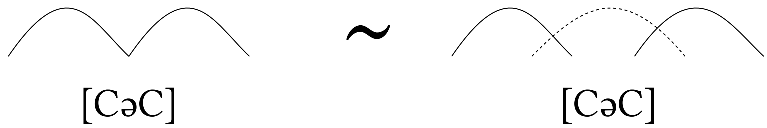


Figure 4: Inconsistent presence of an epenthetic vowel gesture, with open transition (vowel intrusion) in cluster when vowel gesture is absent. Example: Tripolitanian Libyan Arabic (Heselwood et al. 2015): [hat^{ə}k ~hatək].

of Tripolitanian Libyan Arabic. As noted in section 3.1, they find a bimodal distribution of durations for inter-consonantal intervals (ICIs) in lexical consonant clusters, with short ICIs apparently intrusive and longer ICIs apparently epenthetic. Both epenthesis and intrusion occur in phrase-final consonant clusters, as in [hat^{ə}k ~ hatək] 'violation' (transcription inferred from the source). The distribution of ICIs in this specific context shows separate peaks above and below the typical threshold for epenthesis in the dataset. The epenthesis/intrusion distinction can explain why the distribution is specifically bimodal, rather than a wide normal distribution. Note that this level of detail would probably be missed, at least by non-native speakers, without such an acoustic study.

Intrusion/epenthesis variation offers a tool for analyzing certain problematic cases. Another situation where I suspect this might be useful is explaining one anomalous result from Bellik (2019)'s study of intrusion in Turkish CC onsets. As discussed in section 3.3, there is acoustic and articulatory evidence that these vowels are often transitional in speech. Yet when singing such words as part of a text-setting task, speakers who produced the inserted vowel nearly always gave it a note, suggesting it is epenthetic rather than intrusive in that circumstance. One way to reconcile these findings is to consider that speakers might be producing different vowel types in different contexts. Bellik (2019: 179) notes of the singing task that:

> Due to the nature of the task, the non-lexical vowels were produced in a way that sounded impressionistically like a full vowel, in contrast to the brief and schwa-like vowels heard in the acoustic and ultrasound production experiment.

This suggests that some speakers produce intrusive vowels in casual speech, but epenthesize a full vowel gesture as a stylistic flourish when singing. This is somewhat analogous to the P-vocoids that Ridouane & Cooper-Leavitt (2019) posit for Tashlhiyt Berber: epenthetic vowels which occur specifically when needed for realization of a prosodic tune.

Variation between gestural presence and absence is likely to occur at some stage during a diachronic process of reanalysis. Hall (2006) notes a number of cases where vowels that likely began as intrusive are now segmental. But reanalysis does not happen all at once. We would expect a transitional period in which the gestural structures vary, certainly across speakers and possibly within speakers as well. E.g., a single speaker might produce a consonant cluster sometimes with a phonetic transition and other times with an inserted vowel gesture, possibly depending on speech register or prosodic factors. Some innovative speakers might reanalyze all intrusive vowels as segments, at a stage where other speakers still produce them as intrusive.

This possibility of variation between intrusive and segmental vowels obviously complicates the task of describing and classifying vowels. However, recognizing this possibility can help disentangle cases where a vowel insertion pattern seems to be inconsistent.

4.4 Mixed behavior of remnant vowels

Another case where a mix of segmental and nonsegmental vowels might plausibly occur is in what Blevins & Pawley (2010) call the "remnant vowels" of Kalam. Remnant vowels stem from a diachronic reduction process, in which former full vowels reduce to schwa, and then possibly to a purely transitional schwa. Blevins & Pawley (2010) show that that in Kalam, remnant vowels display some properties associated with intrusive vowels and other properties associated with epenthetic vowels[2]. They consider this mixture of characteristics problematic for classifying the vowels as epenthetic or intrusive. However, it is possible that the class of remnant vowels actually encompasses both types.

The vowels in question are also known as "predictable vowels", and occur between all adjacent consonant phonemes, as in /pk-p-n-p/ [ɸɨɣɨβɨnɨp] 'I could have hit' (Blevins & Pawley 2010: 20). Their quality is either central or influenced by nearby consonants or vowels. Speakers do not write them. The vowels appear in contexts where they are not required for repair; they superficially appear to create CV syllables although CVC is licit in the language. All of these characteristics are very typical of transitional, non-segmental sounds.

A couple of other characteristics are less typical of intrusion, but not impossible. The predictable vowels are not speech-rate dependent, and can appear in homorganic clusters. As discussed in section 4.1, these characteristics are predicted to occur when there is a little or no CC overlap (see Figure 2).

[2]Huang (2018) analyzes a similar case in Squliq Atayal (Formosan), which I will not describe here for space, but it poses similar issues.

However, certain predictable vowels show two phonological behaviors that are inconsistent with the phonological invisibility expected of intrusive vowels. First, predictable vowels can appear word-finally to 'bulk up' a subminimal word that would otherwise be just a consonant, as in /m/ [mə] 'taro'. Enforcing word minimality is a phonological repair function typical of epenthetic vowels. Also, predictable vowels bear stress if they are the final vowel in the word, as in /mlp/ [mɨˈlɨp] 'dry'. Both facts are strong evidence for presence of a phonological unit, not merely a transition.

It might be possible to capture both aspects of the vowels' patterning by analyzing predictable vowels as consisting of two distinct groups, similar to Ridouane & Cooper-Leavitt (2019)'s claim that there are both transitional and non-transitional schwas in Tashlhiyt Berber. Although Kalam predictable vowels all result from a diachronic reduction, it could be that reduction was complete in some contexts and incomplete in others. Perhaps in unstressed position, the language experienced full loss of etymological vowels, including loss of their gestures, leaving behind consonant clusters that were phased with very low overlap so as to continue to produce open transitions. In stressed position and in minimal words, however, full deletion was blocked for phonological reasons. Deletion could not occur if the vowel segment was needed to realize stress, or to keep a minimal amount of phonological material in a word. In these contexts, reduction resulted in a vowel that lost its original quality but still had a gesture and formed a syllable. On this analysis, some Kalam predictable vowels are intrusive while others are segmental.

A challenge for this approach is to explain why speakers *don't* write the predictable vowels in cases where they are needed for word minimality, such as /m/ [mə] 'taro'. This does not fit with the assumption that speakers typically write segments, including epenthetic segments. It is possible that such spellings reflect underlying representations, which surface without the predictable vowel in related words like /m-adeŋ/ [ma·nde·ŋ] 'taro plant sp.'

4.5 Symmetrical vowel intrusion? A reconsideration of Hocąk

Hall (2003) argues that Scots Gaelic and Hocąk (Winnebago) have a special type of "symmetrical" vowel intrusion. Both languages have inserted vowels that are relatively long and distinct in quality, yet have a number of characteristics usually associated with intrusion. Briefly, the theoretical proposal is that a sonorant and vowel are adjoined to create a bimoraic nucleus, similar to a diphthong. This is different from most vowel intrusion processes, which typically involve a transition within a complex onset or coda. In this diphthong-like structure, the vowel and sonorant are unordered with respect to one another, and articulatorily realized with their centers aligned. The vowel gesture fully envelops the sonorant, so that periods of the vowel are heard both before and after the sonorant. The overall duration of this bimoraic nucleus is like that of a long vowel.

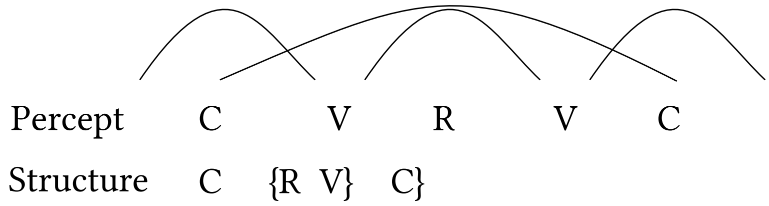


Figure 5: Symmetrical intrusive vowels. C = consonant, R = sonorant, V = vowel. R and V are phonologically unordered with respect to one another.

This proposal attempts to explain why Scots Gaelic and Hoocąk have CVRVC sequences with unusual phonetic properties that pattern with monosyllables in a variety of ways (see Hall 2003, chapters 4 and 5 for details). However, I have come to doubt that this is the correct synchronic structure of inserted vowels in Hoocąk.

For background, Hoocąk inserts copy vowels in underlying /C̥RV/ through "Dorsey's Law", as in /sni/ [sĩnĩ] 'cold'. The resulting [C̥V_iRV_i] structures, known as Dorsey's Law sequences, differ from typical disyllables phonetically and phonologically. Steriade et al. (1990) proposed that the two vowels could be a single gesture, while Alderete (1995) proposed that the sequences are parsed as monosyllables. The symmetrical intrusion analysis ties these ideas.

One problem for this analysis comes from a phonological alternation I had previously overlooked. Susman (1943) reports that the lexical (second) vowel of a Dorsey's Law sequence can undergo raising independently of the inserted vowel. This occurs before suffixes beginning with /-(h)a/, as in /kɾe + hak/ ⟶ [keɾiak] 'he barked moving' (Susman 1943: 75). The second vowel can also undergo independent lengthening before the declarative suffix /-ną/, as in [kiɾiːną] 'he returned' (Susman 1943: 14). In both cases, it looks like the second vowel must be a separate gesture from the first.

Recently, a group in my lab has annotated and begun to analyze archival recordings of a speaker who worked with Kenneth Miner in the 1970's (Hall et al. 2019). We find that Dorsey's Law sequences do have some characteristics that could be consistent with the symmetrical intrusion analysis: they are audibly unlike regular disyllables, with shorter overall duration, and the speaker seems to always produce them as a rapid unit; we have not found cases of pauses or disfluencies within the sequence.

On the other hand, we do not find support for a much-cited claim that the inserted vowels have special accentual properties, which had been important evidence for the intrusion account. Miner (1979) reported that while normal disyllables had one accent, as in [xatáp] 'brush', Dorsey's Law sequences were accented on both vowels, as in [kèré] 'depart returning'. This apparently shared accent seemed like evidence that the inserted vowel formed a unique prosodic unit with the following vowel, compatible with the idea that the two vowels were a single syllable. However, recent work in our lab by Cameron Duval suggests that there is actually no special accent on the inserted vowel. The speaker does produce inserted vowels with a slightly higher pitch than the average lexical vowel in the same position, but the effect is small, and plausibly attributable to the fact that inserted vowels always follow voiceless obstruents, which cross-linguistically raise the pitch of following vowels.

In short, it seems likely that Hoocąk inserted vowels do synchronically have their own gesture, although its short duration indicates that it is phonetically different from lexical vowels. The reanalysis of this example, as well as Hammond et al. (2014)'s challenges to the Scots Gaelic example, make it doubtful that symmetrical vowel intrusion exists.

5 Extending the theory: Gesture-less vowels in other contexts

Although Hall (2006) focused only on intrusive vowels occurring between consonants, the concept of vowel intrusion has also been extended to vowels inserted in other contexts. Various studies have examined the likelihood of non-segmental vocalic periods arising in the transition between a vowel and consonant, or following a word-final consonant, or preceding a word-initial consonant. The evidence for each is reviewed below.

5.1 V-to-C transitions

Several studies have argued that certain gestural timing in VC sequences can produce acoustic intervals that sound like a distinct vowel, without insertion of a gesture. An example is the [a] in Tiberian Hebrew /ruħ/ > [ruaħ] 'spirit'. Operstein (2010) terms this phenomenon "consonant prevocalization", and notes that it is often below the level of speakers' linguistic consciousness, similar to the way that speakers are often unaware of vowel-like releases between consonants. She argues, in an Articulatory Phonology framework, that these vowel-like percepts result from staggered timing of multiple gestures involved in consonants. Each consonant segment, in Operstein's approach, has both a C-gesture and a V-gesture, and sliding the V-gesture earlier produces prevocalization. The effect is the percept of a vowel, which is initially subphonemic. At this stage, the acoustic VVC sequence is in some ways analogous to the acoustic CVC sequences arising from low gestural overlap: there is a percept of an additional vowel, which a linguist can hear but a speaker may not notice, and which does not initially count as an additional syllable in the phonology. It is not truly gestureless, though, since it is produced by a distinct V-gesture (albeit one not associated with a vowel segment).

Others have argued that an intrusive vowel can occur in a V_C context without Operstein's sliding V-gesture, but rather simply due to the tongue's transition between conflicting vowel and consonant articulations. For example, in American English, a schwa-sound often occurs between high tense vowels and liquids, in words like *heel* [hiəl] or *fire* [faɪəɹ]. Gick & Wilson (2006) present evidence that this is an incidental acoustic result of the tongue moving through schwa-like configuration during its transition between conflicting targets. They find that the presence of this schwa does not lengthen the syllable it occurs in. Cohn & Tilsen (2015) find that speakers vary in their syllable count judgments for such words, particularly in the case of /aɪɹ/ and /aɪl/ rimes, as in *fire* and *file*. Words that were judged to have more than one syllable (participants were allowed to give intermediate judgments such as 1.5 syllables) were produced with longer rimes. This might suggest that such words are undergoing reanalysis, with some speakers reinterpreting the transitional schwa as a distinct gesture that forms the nucleus of a new syllable.

Similarly, Garellek (2020) argues that a process in Central Yiddish that inserts schwas between long vowels or diphthongs and certain codas, as in /biːχ/ → [biːəχ] 'book', originated as phonetic transitions. He shows that these schwas occur in contexts where the tongue would be most likely to pass through a schwa-like space when transitioning between the nucleus and coda. Garellek suggests that for some speakers, the schwas are still non-segmental, as shown by their variability. However, some 19th century poetry shows that they counted as metrically present. Like the English schwas discussed above, the Yiddish schwas seem to be prone to reinterpretation as segments.

In short, some vowel-like transitions within V_C do seem to have intrusion-like properties, and also pose similar challenges as to how to interpret gradient or ambiguous speaker intuitions, and mixed or variable patterning.

5.2 Word-final vowel insertion (paragoge)

Could word-final inserted vowels be non-segmental? Among linguists who believe that phonological processes have phonetic precursors, there is widespread agreement that final vowels arise historically as reinterpretation of audible consonant releases (Kang 2003, Blevins 2004:146, 156, Ng 2013). This may explain why, despite being rare as a regular diachronic sound change, final vowel insertion is highly common in second language speech (interlanguage), in loanword adaptation, and in languages that arise from language contact (creoles). Ng (2013) argues that these contexts have in common the slow and effortful speech of adult learners, which is more likely to involve hyper-articulated final consonants with audible releases.

But an origin in audible release does not necessarily mean that there is an intermediate stage in which gestural timing alone produces a stable vowel-like percept, analogous to intrusion. The appearance of the vowel could instead take place in a single leap (released C > CV), with a vowel gesture and segment inserted to satisfy some phonological requirement.

For example, Korean often inserts final vowels into English loans, as in *pad* [p^{h}æ.tɨ]. Kang (2003) points out that this process is not motivated by phonological repair, as codas like [t] are licit in Korean. Her corpus study shows that final vowel insertion is most common in exactly the contexts where English tends to have the most frequent phonetic release, such as final voiced stops. Yet, accepting that the English release is in some sense the source of the inserted vowel, there is no reason to think the inserted vowels are non-syllabic. Kim & Kochetov (2011) show that the insertion is categorical, resulting in a vowel segment whose duration and quality are comparable to those of native vowels. In this case, then, there seems to be a direct leap from the English word with audible (but non-vowel-like) release, to the Korean adaptation which inserts a full vowel in order to maintain a kind of faithfulness to the original release. There is no evidence for a progression of C > Cə > CV, with an intrusion-like stage.

One case where some linguists have argued for non-segmental final vowels is Italian. Consonant-final loanwords are optionally adapted with final schwas, along with lengthening of the final consonant, as in *jet* [dʒɛt ~ dʒɛt:ə]. Miatto (2020) finds that when asked to count syllables in recorded nonce words, Italian speakers judge tokens like [vik:ə] as monosyllabic 93% of the time. When asked to repeat the words, they sometimes insert final [ə] when it was absent from the stimulus, or leave it out when it was present. Repetti (2012) similarly argues that the final vowel "does not play a phonological role". She points to its variable phonetic quality, which she transcribed as [ə], [ᵊ], or [e] (but not identical to lexical [e]), and reports that

> anecdotally, native speakers of Italian do not perceive a final vowel in consonant-final words when they are uttered by themselves or by other native speakers of Italian; however, native speakers of English often do hear a vowel in consonant-final words uttered by Italian native speakers (Repetti 2012: 175).

Grice et al. (2018) mention that there is debate over whether these vowels should appear in the orthography. Grice et al. (2018) argue that Italian final schwa shows a mixture of properties of epenthesis and intrusion. Like epenthesis, it repairs a marked structure, since consonant-final words are rare in the native vocabulary. It is variable, a common characteristic of intrusive vowels, but Grice et al. show that the variability is phonologically conditioned: schwa is more likely to appear on monosyllables than disyllables, and more likely to appear when the word has a complex intonational melody in need of segmental material to realize it. This prosodic conditioning could indicate that the schwa is a phonological element inserted for metrical reasons, similar to the P-vocoid that Ridouane & Cooper-Leavitt (2019) propose for Tashlhiyt (see Section 2). Overall, the question of whether Italian final schwa could be intrusive seems unsettled.

As far as I am aware, there are no proposed articulatory models of final intrusive vowels, where gestural alignment alone can produce a vowel-like percept in final position. One question is how the final period would become voiced, especially if the final C is voiceless as in *jet* [dʒɛtːə]. This is not necessarily problematic, given that Articulatory Phonology treats voicing as a default state. In Browman & Goldstein (1992a), for example, there are no voicing gestures, only wide-glottis gestures to produce periods of voicelessness. In simulations, voicing can appear even in the transition between two voiceless consonants if they are long enough. In the case of final vowel insertion, if the speech system stays "on" after the end of the final oral constriction, the result could conceivably be a short period of voicing that does not correspond to any oral gesture. Whether this actually occurs is an open question. See also Hamann & Miatto (2024 [this volume]) for discussion of the language-specific perception of release bursts.

5.3 Initial vowel insertion (prothesis)

Another context in which vowels are often inserted is word-initially, typically before a CC cluster. Cross-linguistically, prothesis tends to be most common before obstruents and especially sibilants (Fleischhacker 2005). It is unclear whether some initial inserted vowels could lack a gesture. Some impressionistic descriptions seem consistent with this idea; for example, Hewson (1986: 138) notes of Mi'kmaw:

> there is normally a slight prothetic vowel before the initial /k/ in /kti/ 'thy dog' that has traditionally never been written by native speakers, who intuitively recognize it as a purely phonetic element that helps to make the cluster pronounceable.

Operstein (2010) references similar examples from grammars of Welsh and Western Armenian.

Linguists who have explored the possibility of initial intrusive vowels differ in the proposed gestural structure that would underlie such a phenomenon. Bellik (2019: 260) suggests that initial vowel insertion could occur if the nucleus vowel fully overlaps the onset consonant. Operstein (2010: 154–155) views initial vowel insertion as part of the same "prevocalization" phenomenon that also occurs in rimes, where she posits that the "V-gesture" and "C-gesture" of a single coda consonant become temporally disassociated.

Both of these proposals assume that some kind of vocalic constriction gesture must extend before the initial consonant constriction to produce the percept of an initial vowel. Another possible approach, similar to that suggested for final vowel insertion above, would be to explore ways that the default voicing associated with speech in Articulatory Phonology might turn on before the first consonant constriction, producing an initial voiced period unassociated with any oral constriction.

6 Conclusion

The description of vowel insertion processes has benefited enormously from the many careful phonetic and phonological studies reviewed here (as well as others omitted for space). There are now far more acoustic phonetic descriptions and even imaging work on inserted vowels, and more descriptions of how speakers treat the inserted vowels in spontaneous spelling, language games, text-to-tune setting, clapping tasks, etc. Some of this work has supported the intrusion/epenthesis dichotomy, some has challenged it, and all of it lays the groundwork for more detailed and accurate typologies.

This chapter acknowledges that the distinction between vowels with and without gestures is not always as clear-cut as it seemed in Hall (2003, 2006). Characteristics that clustered in a typological survey based mostly on impressionistic descriptions (Figure 1) do not always align perfectly when a single language is probed in depth. I have argued, however, that apparent exceptions can often be explained by 1) recognizing the diversity of gestural configurations that can produce vowel intrusion, 2) exploring the possibility of variation between transitions and gestures, and 3) acknowledging some limitations in our understanding of current methods. Methodological advances in two areas would be particularly useful: more simulations to determine what types of acoustic phenomena can be produced through gestural phrasing alone, and better understanding of what is happening in various metalinguistic judgments and tasks intended to probe speakers' phonological representations.

References

Adiego, Ignasi-Xavier. 2019. Consonant clusters, defective notation of vowels and syllable structure in Caromemphite. In Ronald I. Kim, Jana Mynářová & Peter Pavúk (eds.), *Hrozný and Hittite: The first hundred years* (Culture and History of the Ancient Near East 107), 95–119. Leiden: Brill. DOI: 10.1163/9789004413122_008.

Alderete, John. 1995. Winnebago accent and Dorsey's Law. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Winnebago accent and Dorsey's Law*, 21–51. Amherst: GLSA, University of Massachusetts.

Baronian, Luc & Nicolas Royer-Artuso. 2024. Textsetting the case for epenthesis in Armenian. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 271–289. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264550.

Bellik, Jennifer. 2019. *Vowel intrusion in Turkish onset clusters*. University of California. (Doctoral dissertation).

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette & Andrew Pawley. 2010. Typological implications of Kalam predictable vowels. *Phonology* 27(1). 1–44.

Bosch, Anna & Kenneth De Jong. 1997. The prosody of Barra Gaelic epenthetic vowels. *Studies in the Linguistic Sciences* 27(1). 1–15.

Browman, Catherine P. & Louis Goldstein. 1992a. “Targetless” schwa: An articulatory analysis. In Gerard J. Docherty & D. Robert Ladd (eds.), *Papers in laboratory phonology II: Gesture, segment, prosody*, 26–67. Cambridge, UK: Cambridge University Press.

Browman, Catherine P. & Louis Goldstein. 1992b. Articulatory phonology: An overview. *Phonetica* 49(3-4). 155–180.

Buchwald, Adam, Brenda Rapp & Maureen Stone. 2006. Evidence for discrete phonological representations in production: Ultrasound imaging of aphasic speech. *Brain and Language* 99(1-2). 140–141.

Buckley, Eugene. 2007. Vowel–sonorant metathesis in Alsea. *International Journal of American Linguistics* 73(1). 1–39.

Burke, Mary, Shobhana Lakshmi Chelliah & Melissa Robinson. 2019. Excrescent vowels in Lamkang prefix sequences. *Journal of South Asian Languages and Linguistics* 6(2). 185–213.

Butler, Becky. 2015. Approaching a phonological understanding of the sesquisyllable with phonetic evidence from Khmer and Bunong. In N.J. Enfield und Bernard Comrie (ed.), *Languages of mainland Southeast Asia: The state of the art* (Pacific Linguistics 649), 443–499. Berlin: De Gruyter Mouton.

Canalis, Stefano. 2007. To what extent is Albanian word stress predictable? In Antonietta Bisetto & Francesco Barbieri (eds.), *Proceedings of the XXXIII incontro di grammatica generativa*, 1–14.

Cavirani, Edoardo. In preparation. Epenthetic and intrusive vowels in Lunigiana: How diatopic variation mimics (phonologization) diachrony.

Cohn, Abigail & Sam Tilsen. 2015. Relation between syllable count judgments and durations of English liquid rimes. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow.

Davidson, Lisa & Maureen Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. *Proceedings of the West Coast Conference on Formal Linguistics* 22. 165–178.

Dell, François & Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7. 105–130.

Easterday, Shelece. 2017. *Highly complex syllable structure: A typological study of its phonological characteristics and diachronic development*. University of New Mexico. (Doctoral dissertation).

Fleischhacker, Heidi Anne. 2005. *Similarity in phonology: Evidence from reduplication and loan adaptation*. Los Angeles: University of California.

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

Garellek, Marc. 2020. Phonetics and phonology of schwa insertion in Central Yiddish. *Glossa: a journal of general linguistics* 5(1). 66. DOI: 10.5334/gjgl.1141.

Garmann, Nina Gram, Hanne Gram Simonsen, Pernille Hansen, Holm Elisabeth, POST Brechtje & Elinor Payne. 2021. Cross-linguistic variation in word-initial cluster production in adult and child language: Evidence from English and Norwegian. *Journal of Child Language* 48(1). 1–30.

Gick, Bryan & Ian Wilson. 2006. Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. *Laboratory phonology* 8. 635–659.

Grice, Martine, Michelina Savino & Timo B. Roettger. 2018. Word final schwa is driven by intonation: The case of Bari Italian. *The Journal of the Acoustical Society of America* 143(4). 2474–2486. DOI: 10.1121/1.5030923.

Hall, Nancy. 2003. *Gestures and segments: Vowel intrusion as overlap*. University of Massachusetts Amherst. (Doctoral dissertation).

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2013. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics* 41(2). 133–143.

Hall, Nancy, Andie Niederecker, Elica Sue & Irene Orellana. 2019. Annotating archival recordings of Hocank (Winnebago). In *Proceedings of the Annual Meetings on Phonology*, vol. 6.

Hamann, Silke & Veronica Miatto. 2024. Three language-specific phonological interpretations of release bursts and short vowel-like formants. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 199–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264544.

Hammond, Michael, Natasha Warner, Andréa Davis, Andrew Carnie, Diana Archangeli & Muriel Fisher. 2014. Vowel insertion in Scottish Gaelic. *Phonology* 31(1). 123–153.

Harms, Robert T. 1976. The segmentalization of Finnish 'nonrules'. In *Texas linguistic forum*, vol. 5, 73–88.

Heselwood, Barry, Abdurraouf Shitaw, Aimen Ghummed & Leendert Plug. 2015. Epenthetic and excrescent vowels in stop sequences in Tripolitanian Libyan Arabic. In *Proceedings of the 18th International Congress of Phonetic Sciences*.

Hewson, John. 1986. Syllables and syllabics in Micmac. *Algonquian Papers-Archive* 17. 135–142.

Hickey, Raymond. 1985. The interrelationship of epenthesis and syncope: Evidence from Dutch and Irish. *Lingua* 65(3). 229–249.

Huang, Hui-chuan J. 2018. The nature of pretonic weak vowels in Squliq Atayal. *Oceanic Linguistics* 57(2). 265–288.
Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20(2). 219–273.
Karlin, Robin. 2021. Finnish inserted vowels: A case of phonologized excrescence. *Nordic Journal of Linguistics* 45(1). 49–79.
Kim, Kyumin & Alexei Kochetov. 2011. Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua* 121(3). 511–532.
Kitto, Catherine & Paul V. De Lacy. 1999. Correspondence and epenthetic quality. In C. Kitto & C. Smallwood (eds.), *Proceedings of AFLA (Austronesian Formal Linguistics Association VI*, 181–200. Toronto: University of Toronto, Dept. of Linguistics.
Ladefoged, Peter, Jenny Ladefoged, Alice Turk, Kevin Hind, et al. 1998. Phonetic structures of Scottish Gaelic. *Journal of the International Phonetic Association* 28(1-2). 1–41.
Leonard, Janet. 2007. A preliminary account of stress in Senćoŧen (Saanich/North Straits Salish). *Northwest Journal of Linguistics* 1(4). 1–59.
Levin, Juliette. 1987. Between epenthetic and excrescent vowels. *Proceedings of the West Coast Conference on Formal Linguistics* 6. 187–201.
Miatto, Veronica. 2020. Perception of word-final inserted vowels and syllabicity in Italian. In *Proceedings of the 11th International Conference of Experimental Linguistics (ExLing 2020)*, 137–140.
Michelson, Karin. 1989. Invisibility: Vowels without a timing slot in Mohawk. In Donna Gerdts & Karin Michelson (eds.), *Theoretical perspectives on native American languages*, 38–69. Albany: Suny Press Albany.
Miner, Kenneth L. 1979. Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45(1). 25–33.
Ng, E-Ching. 2013. Paragoge as an indicator of language contact. *Extended Abstracts of the LSA Annual Meeting* 4(31). DOI: 10.3765/exabs.v0i0.794.
Nogita, Akitsugu & Yanan Fan. 2012. Not vowel epenthesis: Mandarin and Japanese ESL learners' production of English consonant clusters. *Working Papers of the Linguistics Circle* 22(1). 1–26.
Operstein, Natalie. 2010. *Consonant structure and prevocalization.* Amsterdam: John Benjamins.
Pariente, Itsik. 2010. Pharyngeal related non-lexical vowels in Sephardic Modern Hebrew. *Linguistics in Amsterdam* 3(1). https://linguisticsinamsterdam.nl/home?issue=142.

Repetti, Lori. 2012. Consonant-final loanwords and epenthetic vowels in Italian. *Catalan Journal of Linguistics* 11. 167–188.

Ridouane, Rachid & Jamison Cooper-Leavitt. 2019. A story of two schwas: A production study from Tashlhiyt. *Phonology* 36(3). 433–456.

Ring, Hiram & James Gruber. 2014. Intrusive vowels in Pnar: There and not there. In *24th meeting of the SouthEast Asian Linguistics Society (SEALS 24)*. Yangon University.

Smith, Kaylin. 2019. Vowel insertion in Scottish English liquid + nasal clusters: Boundary disputes and durational relations. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Vowel insertion in Scottish English liquid + nasal clusters: Boundary disputes and durational relations*, 132–136. Melbourne: International Phonetic Association.

Steriade, Donca et al. 1990. Gestures and autosegments: Comments on Browman and Goldstein's paper. *Between the grammar and the physics of speech: Papers in laboratory phonology* 1. 382–397.

Susman, Amelia. 1943. *The accentual system of Winnebago*. Columbia University. (Doctoral dissertation).

Tabain, Marija & Birgit Hellwig. 2022. Qaqet. *Journal of the International Phonetic Association* 53(3). 1073–1094.

Van Donselaar, Wilma, Cecile Kuijpers & Anne Cutler. 1999. Facilitatory effects of vowel epenthesis on word processing in Dutch. *Journal of memory and language* 41(1). 59–77.

Walter, Mary Ann. 2018. The significance of vowel insertion in Ottoman Turkish. In *The Rouen Meeting: Studies on Turkic Structures and Language Contacts*, 1st edn., 87–94. Harrassowitz Verlag.

Warner, Natasha, Allard Jongman, Anne Cutler & Doris Mücke. 2001. The phonological status of Dutch epenthetic schwa. *Phonology* 18(3). 387–420.

Wiik, Kalevi. 1965. *Finnish and English vowels: A comparison with special reference to the learning problems met by native speakers of Finnish learning English*, vol. 94. Turun: Turun yliopisto.

Chapter 9

Three language-specific phonological interpretations of release bursts and short vowel-like formants

Silke Hamann[a] & Veronica Miatto[b]
[a]ACLC, University of Amsterdam [b]Stony Brook University

In this paper we challenge the assumption that there is one correct interpretation of auditory signals by showing that the mapping between perceptual cues and phonological categories is highly language specific. This is exemplified by three different languages and their different interpretations of final release bursts.

The first case is American English (AE), where a release burst in final position is always interpreted as a perceptual cue to a final plosive by AE listeners. The second case is the interpretation of a final release burst as a vowel, despite the lack of vowel-like formants following the burst, as in Korean. In the third case, exemplified by Italian, not only a release burst but also vowel-like formants are present word-finally, but the latter are not interpreted as phonological vowels.

These different phonological interpretations of the same phonetic material can only be formalized in a linguistic model that makes a clear distinction between phonetic and phonological representations, and allows for language-specific, non-universal mappings between the two. For this, we employ the OT version of the Bidirectional Phonetics and Phonology model, and show how different rankings of cue constraints in these three languages can account for the different phonological interpretations. This formalization accounts for both native and naive second language perception, and provides testable predictions for future experimental studies.

Silke Hamann & Veronica Miatto. 2024. Three language-specific phonological interpretations of release bursts and short vowel-like formants. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 199–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264544

1 Introduction

Speech perception involves the interpretation of auditory information as phonological units, and is to a certain degree language specific, as early illustrations by e.g. Polivanov (1931/1974) have already shown. A familiar example of this language specificity is the phenomenon of so-called "illusory" vowels: listeners exposed to words with consonantal clusters that are illicit in their native language perceive a vowel breaking up these clusters, even though the acoustic signal does not contain any corresponding vowel-like formants, see e.g. Dupoux et al. (1999) for Japanese, Kabak & Idsardi (2007) and Durvasula & Kahng (2015) for Korean. Interpretations of this phenomenon mainly focus on the role of language-specific phonotactic restrictions and the absence of vowel formants, the latter leading to its name. Less attention is paid to the fact that these perceptual "illusions" seem to be an efficient way for listeners to interpret their native language, where the corresponding vowel categories are often realized as voiceless: In Japanese, for instance, the vowel /ɯ/ is usually devoiced if preceded and followed by voiceless consonants, therefore it helps Japanese listeners to perceive the absence of vowel formants in such a context as a vowel /ɯ/ to quickly retrieve the correct word form (see e.g. the results of the lexical decision task with Japanese listeners by Ogasawara & Warner 2009). In speech perception, language-specificity thus holds not only for phonotactic restrictions, but also for the interpretation of perceptual cues.

The assumption of a language-specific mapping between auditory information and phonological categories opposes the view of those generative phonologists that assume a universal phonetics-phonology interface, see e.g. Hale & Kissock (2007) and Hale & Reiss (2000) (cf. Hamann 2011 for a detailed discussion). In their view, sounds that share the same phonological feature (or feature value) across languages have identical auditory cues and/or articulatory gestures.

In the present article, we illustrate with the example of two possible cues for plosives, namely burst release and a following short period of vowel-like formants, how language-specific the interpretation of perceptual cues can be. There are, of course, many more language-specific plosive cues, such as voice-onset time (see, e.g., the seminal papers by Lisker & Abramson 1964, 1970), which we do not consider here for reasons of space and clarity. We restrict our discussion to velar plosives in three languages, American English, Korean and Italian, as summarized in Table 1. With respect to Italian, we focus on the Venetian variety and ignore variation in the interpretation of final schwa-like formants that have been reported for several other dialects (see e.g. Cavirani 2015 for the Lunigiana dialects).

Table 1: Cues for syllable-final velar plosives in three example languages

Cue interpretations	Example language
Closure release is an optional cue to a syllable-final plosive Short period of formants following the release is an obligatory cue to a following vowel	American English
Closure release is an obligatory cue to an onset plosive Short period of formants is a cue to a following vowel	Korean
Closure release is an obligatory cue to a syllable-final plosive Short[a] period of formants is an optional cue to a word-final plosive	Italian

[a]This period shows quite some variability in duration, as mentioned in Section 2.3. Nevertheless it is on average shorter than a full vowel.

The pattern of often-released coda plosives in American English can also be observed in the closely-related West-Germanic languages Dutch and German, and in Romance languages (where the coda release seems obligatory, see §2.3 below). The dominance of researchers with a West-Germanic or Romance native language in the linguistic field from the 16th until the mid 20th century, but also extensive, detailed phonetic studies of these restricted language groups might have led to the false assumptions that 1) the release of coda plosives is a common pattern in all languages, and 2) that such a release has to be interpreted as indicating a plosive, only. A lack of a plosive burst in syllable-final position, however, is obligatory in our second example language, Korean, and also in other languages spoken in Asia, such as Hong Kong Cantonese (Bauer & Benedict 1997), Thai (Abramson & Tingsabadh 1999), Vietnamese (Kirby 2011), and elsewhere, e.g., Karitiana (Storto & Demolin 2002), Ibibio (Urua 2004), and Efik (Cook 1969). Korean listeners interpret a coda burst as indicating a following (voiceless) vowel. Formant-like structures, on the other hand, are interpreted in many languages as indicating a vowel category, while in our third example language, Italian, they are a cue enhancing the release of a final plosive. A similar interpretation of short, vowel-like formant structures after coda consonants can be found in French (see, e.g., Flege & Hillenbrand 1987).

Besides a description of how the three example languages given in Table 1 interpret these two auditory cues, the present article also provides an explicit formalization, i.e. the mapping of these cues onto phonological categories (such as plosives and vowels). If this mapping were universal, language learners would not need to acquire it, and linguists would not need to model it. However, since we observe language-specific differences in the interpretation of these cues, we provide a model of these language-specific mappings to account for the knowledge that the learners of these three languages have acquired. To perform such a formalization, a model is needed that makes a strict distinction between phonetic realizations and phonological categories, and that allows for a language-specific mapping between the two. For this purpose, we employ Bidirectional Phonetics and Phonology (Boersma 2007, 2011, 2009), the only existing linguistic model to our knowledge that incorporates detailed phonetic realizations as well as phonological surface forms (categories but also other phonological structure like syllables, prosodic words, etc.). We will use this model to illustrate that by formalizing the details of L1 perception, i.e., which cues are mapped onto which phonological form, we can account for what listeners of these languages do when encountering an auditory form that does not exist in their L1 (i.e., we model the process of naive L2 perception).

This article is structured as follows. We briefly describe the relevant phonetic and phonological details of the three languages American English, Korean and Italian in §2. §3 provides the formalization, and §4 then discusses and concludes.

A short explanation of the notation we use is necessary. In order to distinguish three levels of representations, the present article employs square brackets for [auditory forms], slashes for /phonological surface forms/, and pipes for |underlying, lexical forms|.

2 The three example languages

2.1 American English

American English (henceforth: AE) is a language that has plosives both in onset and in coda position, as e.g. in *cake* |kʰeɪk|. In both onset and coda position, plosives can be followed and/or preceded by other consonants. Syllable-final coda plosives are often released with an audible burst, but can also be unreleased (e.g., Rositzke 1943, Crystal & House 1988, Davidson 2011). According to Kim (1998), this is a notable difference with Korean, where a plosive in coda position is never released, see elaboration in §2.2 below.

Whether coda consonants in AE have a burst cue depends on many factors: Plosives are more often released after tense vowels than after lax vowels (e.g., Parker & Walsh 1981), in phrase-final position plosives than in phrase-medial position (e.g., Crystal & House 1988), by female speakers than by male speakers (e.g., Byrd 1993), and velar plosives are more often released than coronal ones (e.g., Halle et al. 1957, Crystal & House 1988, Byrd 1993), to mention only the most often observed factors.

Though the presence of a burst cue is not necessary for AE listeners to perceive a coda plosive, its absence can be falsely interpreted as a syllable without coda plosive, as indicated by several perception studies (e.g., Householder 1956; Halle et al. 1957, who show that this is particularly the case for plosives following lax stops; and Lisker 1999, who shows that this holds for plosives, especially velars, following diphthongs).

AE has not been reported to employ short vowel-like formants after the release burst to enhance the burst. AE listeners perceive speakers of languages like Italian that use such vowel-like formants following a plosive release as producing a full vowel in this position (e.g., Hall 2006), and an Italian accent in English is therefore often caricatured with schwas or [e]-like vowels after each final consonant (see also Busà 2008).

2.2 Korean

Korean speakers do not release their syllable-final plosives (e.g., Martin 1951, Kim 1998, Kang 2003). Underlyingly aspirated plosives in this position also lose their aspiration, see, e.g., the neutralization of underlying |puʌk^{h}| 'kitchen', which is realized as surface /.pu.ʌk./, compared to /.pu.ʌ.k^{h}ɛ./ 'in the kitchen' (Kang 2003: 224; notation adapted).

Due to the absence of bursts in syllable-final position, Korean listeners interpret a plosive release in the auditory signal as an indication that the plosive occurred in syllable-initial position and was followed by a vowel, even if there are no formants in the auditory signal to support the percept of a vowel. Since the high vowel /ɨ/ in open syllables is often devoiced before and after voiceless obstruents in Korean (Kim-Renaud 1987, cited by Kang 2003: 236), Korean listeners regularly interpret consonant clusters that are phonotactically illegal in Korean as being broken up by a voiceless vowel /ɨ/ (Kang 2003; see Durvasula et al. 2018 for a detailed account of the choice of vowel).

The effect of this perceptual interpretation can be extensively observed in vowel insertion both in the adaptation of loanwords and in experimental studies, as a large body of literature has shown (e.g., Durvasula & Kahng 2015, 2016, Hutin 2014, Kang 1996, Kang 2003, Kabak & Idsardi 2007, Iverson & Lee 2006). While for velar plosives the perception with a following vowel is almost categorical, see e.g. English *spike*, which is borrowed into Korean as /.sɨ.pʰɑ.i.kʰi./, this is more variable for other places of articulation, see e.g. English *flute*, which is borrowed both as /.pʰɨl.lu.tʰi./ and /.pʰɨl.lut./[1] (examples from Kang 2003; notation adapted). The vowel insertion in such cases depends on many factors, amongst them the tenseness of the preceding vowel, the fact that Korean nouns show an alternation between /s/ and /t/ in final position, the actual realization of these words with a final release in the donor language (such as AE, cf. §2.1 above), the functional load of the word, etc. (see, e.g., Kang 2003, Chang 2018, Kim 2022, for extensive discussions). We will come back to this variation in §4.2.

2.3 Italian

In Italian, word-final consonants are always released, and are sometimes produced with vowel-like formants after the release, so a word like *jet* might be pronounced as [dʒɛtə] and a word like *tunnel* might be pronounced as [tunnelə] with lengthening of the last consonant. The presence of these word-final, schwa-like formant structures varies based both on inter-speaker (exposure to English) and intra-speaker variables (the type of word-final consonant, the number of repetition etc.), as explained below. In production, consonant-final words *can* therefore be followed by schwa-like formant structures, but for the majority of speakers this is not categorical (see Miatto et al. 2019, Miatto 2022). However, in perception, which is the focus of this paper, a consonant-final word like *jet* will be categorically perceived as consonant-final whether the consonant is followed by such vowel-like formants or not (see Miatto 2020 for a relevant study using nonce words).

Italian consonant-final words are relatively new and are mainly loanwords or acronyms, so these vowel-like formant structures have long been treated as vowels epenthesized to adhere to a phonological constraint that prohibits word-final codas (Bafile 2002, 2003, 2005 for Tuscan Italian, Passino 2008 for Abruzzese Italian, and Broniś 2016 for Roman Italian). The implications of this interpretation are that 1) the phonetic material is perceived and produced as a phonological vowel, and that 2) it constitutes the nucleus of a separate syllable. Although using different frameworks, these authors agree that *jet* in Italian is constituted by two syllables /dʒɛt.tə/, with the lengthened consonant serving as a geminated consonant.

[1] For this particular word, the borrowing with a coda consonant, thus without inserted vowel, occurs more frequently, as a Google search by a reviewer showed.

More recently, however, it has been proposed that these vowel-like elements are not syllabic, i.e., they do not constitute the nucleus of a syllable. Grice et al. (2018) propose that for Barese Italian it is a non-syllabic vowel that appears under certain phonological and prosodic pressures, and its characteristics are similar to Hall's (2006, 2011) intrusive vowels (see also Hall 2024 [this volume]). Similarly, Repetti (2012), followed by Miatto et al. (2019) and Miatto (2020, 2022) for Venetian Italian, propose that this is a vowel-like segment that is part of the release of the consonant. The latter is also the view we adopted here, and we will refer to the presence of these vowel-like formant structures after the release of a word-final consonant as "excrescent vowel". Based on studies on Venetian Italian, we are following the latter theoretical interpretation of excrescent vowels for the following reasons. Note that from here on, when we refer to "Italian" we specifically refer to the variety of Italian spoken in Veneto (North-East Italy).

First of all, the occurrence rate of excrescent vowels in Italian is susceptible to experience with English, which does not display such vowel-like formant structures after coda releases (recall §2.1). Miatto et al. (2019) found that with increasing experience in spoken L2 English, speakers tended to produce fewer excrescent vowels after word-final consonants in Italian. This suggests that excrescent vowels are likely to be a phonetic phenomenon rather than a phonological repair, because the speakers' exposure to L2 English was probably too limited to have caused a change in their L1 phonology.

Second, following Hall (2006, 2011)'s diagnosis of epenthetic (i.e., phonological) vowels versus intrusive vowels (formant transitions between consonants), the characteristics of excrescent vowels in Italian are much closer to intrusive vowels than canonical epenthetic vowels. Excrescent vowels in Italian are not always present, they acoustically resemble schwa, and are highly variable in their duration. Moreover, they do not participate in certain phonological processes such as stress assignment (Repetti 2012) or syllabicity (Miatto 2020). In particular, Miatto's (2020) study on the perception of word-final excrescent vowels shows that Veneto speakers are not aware that they insert an excrescent vowel, and they do not perceive it. The participants of the study listened to nonce words such as /vik/, to which excrescent vowels of varying durations (incremental steps of 25 ms ranging from 0 to 100 ms) were inserted after the word-final consonant, and judged them as monosyllabic 93% of the time. Moreover, the duration of the word-final schwa did not influence significantly whether they would perceive the nonce word as a monosyllable or disyllable.

Third, Miatto's (2022) findings on factors that condition the presence of excrescent vowels after voiceless plosives can only be explained by referring to their perceptibility. In her study, excrescent vowels were more likely to appear after labials and coronals than dorsals. It was argued that in Italian, since the overwhelming majority of consonants is followed by a vowel, formant transitions are extremely important cues for the perceptibility of plosives. Word-final labials and coronals, which have weaker bursts (Dorman et al. 1977), would then be less perceptible than dorsals, unless you had plosive releases that incorporated formant transitions. Another finding of Miatto's (2022) study is that with increased repetitions of the same plosive-final nonce word, the presence of excrescent vowels significantly decreased. Repetition is generally shown to have a negative effect on clarity and intelligibility of phonetic production (Fowler & Housum 1987), supporting the argument that the Italian excrescent vowel is a phonetic cue that aids the perceptibility of the final consonant, but appears less if the words are produced less carefully in later repetitions.

Finally, as stated in Miatto (2022), duration measurements were not consistent with a phonological analysis in which the excrescent vowel is syllabic and the nucleus of a separate syllable. In the study, she found that vowels in nonce words such as *fap* were short, and therefore obligatorily in a closed syllable due to their duration. Moreover, word-final consonants in nonce words were significantly shorter than control geminated consonants, which indicates that the consonants might not be geminated but only slightly lengthened, contrary to what was reported in previous literature.

3 Modelling the language-specific perception of burst release and vowel formants

In this section, we formalize the language-specific use of the plosive cues described in the previous section, making explicit the knowledge that the listeners of the three languages employ when listening to their native input. We also apply this knowledge in the form of a native perception grammar naively to input from the other two languages unknown to the listeners (naive L2 perception), and partly compare this to reported results from L2 perception and loanword adaptation (assuming adaptation took place via L2 perception, though alternative adaptations by bilinguals are also possible, see e.g. Paradis & LaCharité 1997).

As mentioned in the introduction, we employ Bidirectional Phonetics and Phonology (henceforth: BiPhon; Boersma 2007, 2011, Boersma & Hamann 2009) for our modelling, because it provides an explicit formalization of the phonetics-phonology interface by mapping auditory cues onto phonological surface categories, and vice versa. The relevant representations and constraint types to formalize this mapping are given in Figure 1.[2]

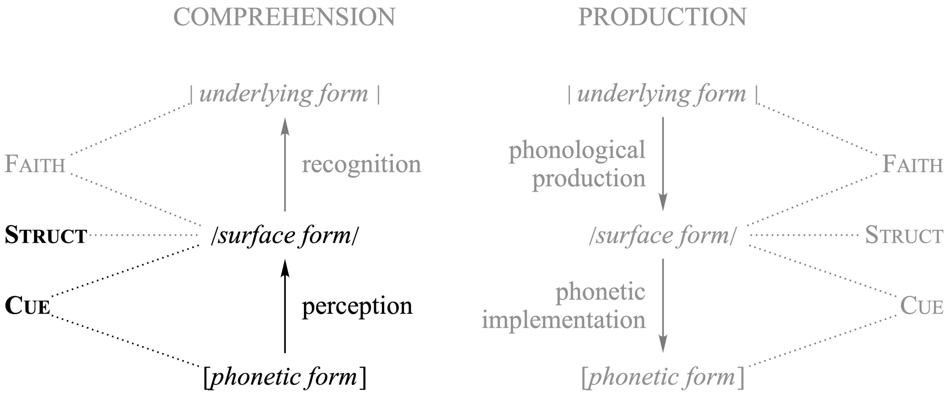


Figure 1: The BiPhon model, with representations in italics and constraints in small capitals. The perception process is given in black. This mapping from auditory to surface representation and its reverse mapping in phonetic implementation form the phonetics-phonology interface.

The representations and mappings in BiPhon correspond to those in psycholinguistic models of comprehension (e.g., McQueen & Cutler 1997) and production (e.g., Levelt 1989). As the present article is only concerned with (native and non-native) speech perception, the formalization is restricted to the perception process, any influences of lexical forms in speech comprehension are ignored, but are of course relevant in the perception of real words.[3]

The perception grammar consists of two types of constraints: Cue constraints, mapping the auditory onto the surface form, and Struct(ural) constraints, restricting the surface form. These constraints and their rankings are also used to model the production direction, see e.g. Boersma & Hamann (2009) for an illustration. The perception grammar is thus no additional device modelling only speech perception, but an integral part of a listener's/speaker's grammar. Our formalization is restricted to the voiceless velar plosive /k/ and a preceding front

[2] Note that Figure 1 makes no distinction between auditory and articulatory form, and summarizes both under [phonetic form]. As the present article is only about speech perception, this distinction is not relevant, but the interested reader is referred to Hamann (2011) and the discussion therein on the precise order of these forms in BiPhon and a comparison to alternative grammar models.

[3] See, e.g., Boersma (2009) for the modelling of the Ganong effect in BiPhon with parallel evaluations of surface and underlying forms.

mid vowel /ɛ/, both occurring in all three languages. The formant cues corresponding to the vowel are summarized by the auditory form [ɛ], as the detailed cues are not of relevance in the present article, but even though the same symbol is used for the auditory as for the phonological form, the reader needs to keep in mind that [ɛ] represents values of the first three formants, duration, and other auditory information of a typical realization of the abstract phonological category /ɛ/. The auditory cues of the velar /k/ of relevance to our formalization are the following: a closure preceded by the vowel, represented as the sequence [ɛ ˺_], where [˺] stands for the vowel transitions containing information on the velar place of the following plosive, and [_] for the silent closure during the voiceless plosive. Furthermore, there is the velar release burst, represented as [ᵏ]. In addition, we consider the excrescent vowel that can be found in Italian final plosives after the release of the burst, and will notate these vowel-like formants as [ᵊ], to represent their often very short duration. An inclusion of other cues and other places of articulation than velar would go beyond the scope of this article.

In the following sections, we formalize the language-specific interpretation of these auditory cues, starting with the minimal cues of an unreleased [ɛ ˺_], and continuing by adding subsequently the burst and vowel-like formant cues to the auditory input of the perception tableaux. We will show how these inputs, even though not always native to the language in question, are dealt with by the native perception grammars of the three languages.

In the assumption that the raw auditory signal is perceived as an abstract category, and not stored as such, we depart from theories like Exemplar Theory (Pierrehumbert 2001), that presume listeners have a holistic memory trace of all the acoustic details of an auditory input that they encountered.

The constraint rankings that we employ in the following is strict, and therefore result in categorical behaviour, i.e., one candidate wins. It is well-known that humans do not exhibit such categorical behaviour, and that the percept depends on several linguistic and social factors. We will elaborate on how this variation can be integrated in our model in §4.2 below.

3.1 Perception of transitions into closure and the closure phase

The auditory form [ɛ ˺_] is perceived as the phonological surface /ɛk/ due to the transitions into the closure and the silence during closure.[4] This is captured by

[4]A reviewer voiced scepticism about listeners' ability to perceive silence. Silence, i.e., the absence of periodicity or friction noise, and its relative duration, has been shown to be picked up by listeners and to be an important cue for distinguishing plosives from fricatives and affricates (e.g., Repp et al. 1978, Dorman et al. 1979).

Cue constraints like *[˺_]/s/, which stands for "Do not map velar transitions and a silent closure onto an alveolar fricative in the surface form". Cue constraints are employed in BiPhon to map auditory information onto phonological categories, and are formalized negatively due to OT's exclusion mechanism.[5] Further, similar Cue constraints could be employed to exclude other consonantal candidates but we refrain from this in the interest of brevity and clarity. An antagonistic constraint, violated when these cues are mapped onto a phonological velar plosive, i.e., *[˺_]/k/, though seeming counter-intuitive, is relevant and included in the following formalization. The constraint *[˺_]/ / avoids that these cues are simply ignored (mapped onto nothing in the surface form). Table 2 illustrates how these three constraints capture the correct mapping onto a velar voiceless plosive in Korean.

Table 2: Korean

[ɛ˺_]	*[˺_]/s/	*[˺_]/ /	*[˺_]/k/
☞ a. /ɛk/			*
b. /ɛs/	*!		
c. /ɛ/		*!	

In the perception tableaux in Table 2, as in the following perception tableaux, the input is an auditory form, and the output candidates are surface phonological forms, which are constructed by the listener to access underlying phonological representations in the mental lexicon (i.e., the intermediate stage in psycholinguistic models of speech perception and comprehension by, e.g., McQueen & Cutler 1997). The input is a native production of a Korean unreleased coda plosive.

Under ideal circumstances (without background noise), Korean listeners perceive [ɛ ˺_] as surface /ɛk/, i.e., as a velar plosive in coda position, because they are used to non-released plosives in their language and familiar with auditory inputs like these to be interpreted as surface representations of coda plosives. We therefore assume that the two first constraints in Table 2 are high ranked and the third low ranked in the perception grammar of Korean. AE listeners are also used

[5]With a set of Cue constraints that prohibits the mapping of all possible occurring auditory events onto all possible phonological categories, and input distributions of actually occurring auditory cue values for phonological categories in a language, one can simulate the acquisition of such a perception grammar and thus provide a stochastic model of the language-specific acquisition process of a child (see, e.g., Escudero & Boersma 2003). A more realistic model would assume positive connections between occurring values and phonological categories (as possible in Harmonic Grammar (Legendre et al. 1990), see, e.g., Zhou & Hamann 2024).

to perceiving unreleased plosives as coda consonants, and we therefore assume a similar ranking for now, though we learned in §2.1 that the absence of a burst can also be interpreted by AE listeners as a syllable without a coda plosive. We will return to this variability in AE perception in §4.2.

For Italian, unreleased plosives in coda position are not reported (recall §2.3). We therefore assume Italian listeners without any knowledge of languages like English or Korean are likely to perceive the non-native [ɛ ˺_] as not containing a plosive. They require explicit cues for the existence of a plosive consonant, as we will see in the following sections. Rather, we speculate that Italians perceive this input as containing only a vowel, i.e., as the third candidate in Table 3. To the best of our knowledge, experimental evidence for these assumptions do not exist, and a future perception study will need to test them. However, our speculations are based on impressionistic evidence, namely the second author of this paper noticing that some native Italian participants were occasionally perceiving experimental nonsensical stimuli like [mip] as /mi/ when asked to repeat what they were hearing. It has to be noted that those particular stimuli had a very quiet bursts and no word-final schwa-like formants. This leads us to speculate that without a burst (or even with a very quiet one, as represented by these cases) Italian speakers do not perceive a final plosive consonant.

Based on our assumptions, we argue that in the Italian perception grammar, the third candidate wins, and therefore the constraints *[˺_] / / and *[˺_]/k/ have to have a reverse order than in Korean, see Table 3 for the Italian perception grammar:

Table 3: Italian

[ɛ˺_]	*[˺_]/s/	*[˺_]/k/	*[˺_]/ /
a. /ɛk/		*!	
b. /ɛs/	*!		
☞ c. /ɛ/			*

In Sections 3.2 and 3.3 we will not include these three constraints, nor the candidates two and three, but we will come back to them in §3.4.

3.2 Perception of a burst release

We continue by adding the cue of a burst release, which natively is used in AE for coda plosives. We thus consider how our three groups of native listeners would cope with a typical AE input (native perception for AE listeners, naive L2 perception for Italian and Korean listeners).

The addition of the burst to the auditory input, resulting in the input [ɛ ̚ _ᵏ] requires two separate Cue constraints. The first, *[ᵏ]/ /, is violated when the burst cue is ignored. This is not allowed in any of the three languages; hence the constraint must be highly ranked in all three. The second constraint avoids that the burst is interpreted as a coda consonant: *[ᵏ]/k./, where "." stands for a syllable boundary. In Korean, where the burst is an indication that the consonant occurred in onset position, this constraint is high ranked, in AE and Italian low.

We learned in §2.2 that in cases where there are no vowel cues, Koreans perceive an /ɨ/, as this vowel is often devoiced. The input [ɛ ̚ _ᵏ] would thus be mapped onto the phonological form /ɛ.kɨ/ by naive Korean listeners, a mapping that violates the Cue constraint *[ ]/ɨ/: "Don't map the absence of vowel cues onto a surface /ɨ/". This constraint is low ranked in Korean, see Table 4, due to the often-occurring devoicing, i.e., listeners are used to interpret a non-existence of formants in the auditory form as the vowel category /ɨ/ in Korean.

Table 4: Korean

[ɛ̚_ ᵏ]	*[ᵏ]/ /	*[ᵏ]/k./	*[ ]/ɨ/
a. /ɛk/		*!	
☞ b. /ɛ.kɨ/			*
c. /ɛ/	*!		

For AE, we assume that the vowel that is closest to having no perceptual cues is the unstressed schwa, and hence the relevant candidate and constraint look slightly different, cf. Table 5. In contrast to Korean, AE does not allow listeners to perceive a vowel (even a schwa) in the absence of corresponding cues, as reflected in the high ranking of the constraint *[ ]/ə/.

Table 5: AE

[ɛ̚_ ᵏ]	*[ᵏ]/ /	*[ ]/ə/	*[ᵏ]/k./
☞ a. /ɛk/			*
b. /ɛ.kə/		*!	
c. /ɛ/	*!		

Italian has similar candidates and a similar constraint ranking as AE in Table 5, since a velar release is in general very strong (Dorman et al. 1977) and therefore a sufficient cue for a plosive, also in Italian.

3.3 Perception of short vowel-like formants following the burst

The last cue that we include in our formalization is the excrescent vowel that is typical in the production of coda consonants in Italian, resulting in the auditory input [ɛ˺_^{kə}]. We will illustrate how Italian listeners perceive this input natively, and how the other two language groups deal with this non-native input.

Again, this cue requires constraints referring to it. The constraint *[^{ə}]/ / is taking care that this cue is mapped onto a separate phonological surface vowel, and is high ranked in AE and Korean. In Italian, it is low ranked, and here the short vowel-like cue is interpreted together with the burst to indicate a velar occurring in final position; it enhances the burst cue. We can formalize this with the constraint *[^{kə}]/.k/, which is high-ranked in Italian, see Table 6.[6]

Table 6: Italian

[ɛ˺_^{kə}]	*[^{k}]/ /	*[ ]/ə/	*[^{kə}]/.k/	*[^{k}]/k./	[^{ə}]/ /
☞ a. /ɛk/				*	*
b. /ɛ.kə/			*!		
c. /ɛ/	*!				

Interestingly, this native Italian perception grammar predicts that naive Italian listeners without experience of a foreign language would perceive words in a foreign language that have final unstressed schwa categories as having no schwa (or any other vowel) in final position. In production, the Italian listeners would then realize such words with schwa-like formant structures due to the use of the same Cue constraint in the production direction (cf. Figure 1). AE listeners, on the other hand, perceive short vowel-like formants as a full vowel, see the respective constraint ranking in Table 7.

A similar ranking of *[^{ə}]/ / above *[^{kə}]/.k/ in the Korean perception grammar accounts for the fact that Korean native listeners also perceive such a schwa as separate vowel.

[6]Non-velar plosives have in general weaker bursts (Dorman et al. 1977), and for them the cue of burst alone might not be sufficient to be perceived as a plosive in Italian. This could be modelled e.g. for an alveolar plosive with the following Cue constraints and their ranking:

*[^{tə}]/.t/ ≫ *[^{t}]/t./ ≫ *[^{t}]/ /

Table 7: AE

[ɛ˺_ ^{kə}]	*[^{k}]/ /	*[ ]/ə/	*[^{ə}]/ /	*[^{k}]/k./	[^{kə}]/.k/
a. /ɛk/			*!	*	
☞ b. /ɛ.kə/					*
c. /ɛ/	*!				

3.4 Summary of the constraint rankings

In Figures 2–4 the constraint rankings of the three languages are summarized in Hasse diagrams. In these diagrams, constraints in the upper stratum are higher ranked than those in the lower stratum to which they are directly connected by a line. For the three constraint sets that are not connected via lines, the ranking between these sets cannot be established (this holds for all three languages).

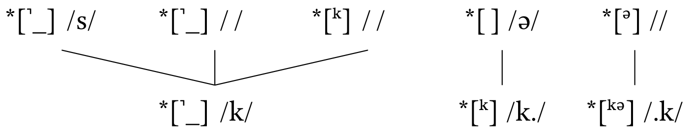


Figure 2: Hasse diagram of the constraint rankings in AE.

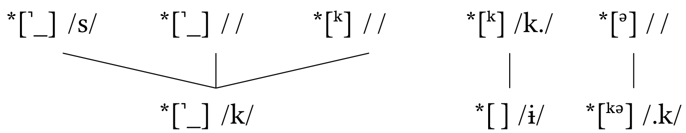


Figure 3: Hasse diagram of the constraint rankings in Korean.

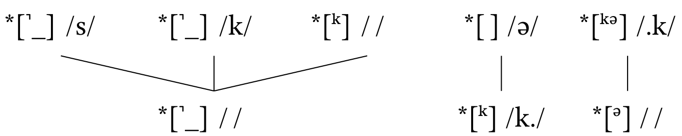


Figure 4: Hasse diagram of the constraint rankings in Italian.

As we can see, Korean (Figure 3) differs from AE (Figure 2) in the order of the two constraints responsible for the interpretation of the burst as coda plosive and of the silence as vowel (AE /ə/ and Korean /ɨ/), cf. the second, middle set of constraints. Italian (Figure 4) differs from AE in the order of the constraints for the interpretation of the burst with short formants as onset plosive and of the short formants as nothing, cf. the third, right-most set of constraints, and from both AE and Korean in the order of the two constraints against interpreting formant transitions and silence cues as velar plosive or as nothing, cf. the middle of the first, left-most set of constraints.

None of the three language-specific interpretations of the two cues involved any Struct constraints, i.e. is caused by restrictions on the syllable structure in the respective languages, though the reader needs to keep in mind that phonotactic restrictions such as on possible Coda consonants could play a role in perception and are therefore part of the perception grammar, where they evaluate the phonological surface forms (see Boersma & Hamann 2009 for an illustration from Korean).[7]

Though we have treated here the set of constraints as universal in order to facilitate a cross-linguistic comparison, we assume that the cues and the phonological categories used in a language are acquired on the basis of the input that the child receives (thus through statistical inference), and are not innate.

4 Discussion and conclusion

In §3 we illustrated how listeners of the three languages American English, Italian and Korean interpret the auditory cues of burst noise and a short period of vowel-like formants differently, and how this can be formalized by means of three perception grammars that differ in the (rankings of) Cue constraints. Only a grammar model that makes a systematic distinction between phonetic/auditory form and phonological representations, and which allows for a language-specific mapping between the two, such as BiPhon, can provide such a formalisation. Grammar models without such a distinction (such as, e.g., the two-level OT models by Flemming 2001 and Steriade 2001) need to introduce extra-grammatical devices (such as Steriade's p-map) to refer to possible auditory cues, while a universal mapping (as in the three-level models by e.g. Hale & Kissock 2007 and Hale & Reiss 2000) would not allow for any differences between languages.

[7] A reviewer wondered whether a Cue constraint like *[$^{\text{kə}}$]/.k/ does not incorporate structural information as it refers to the syllable boundary. Cue constraints like these express the fact that certain cues only occur in certain positions, but are inherently different from Struct constraints like NoCoda, which are cue-independent restrictions on the phonotactic structure of a language.

4.1 Why an explicit formalization?

A reviewer asked what a formalization like we performed in §3 can buy us. This is a valid question we would like to answer in this subsection. Any kind of linguistic formalization makes explicit the knowledge a speaker/listener has acquired about their language (in line with the general aim of a linguistic model). In contrast to a simple lists of this knowledge, an explicit formalization makes use of a restricted set of tools (such as, e.g., the Cue and Struct constraints in BiPhon, and the exclusion mechanism of OT), which force the scientist to also consider other logical possibilities (e.g., opposite Cue constraints or alternative candidates) that are often ignored when thinking in terms of simple lists, but which are necessary to fully capture all relevant information, e.g. language-specific knowledge.

Furthermore, an explicit formalization is able to make predictions. Our formalization of the Italian perception grammar and its application to non-Italian input, e.g., resulted in the prediction that naive Italian listeners without experience with a foreign language will perceive input that contains a final unstressed schwa in the surface form of the foreign language, and thus schwa-like formant structures in the auditory form, as having no schwa (or any other vowel) in final position. This prediction needs to be tested with perception experiments, and the experimental results can then inform us about the correctness of our assumptions. If a post-plosive final schwa in e.g. English will be perceived as separate phonological unit, then we were incorrect in assuming that schwa-like formant structures only function as enhancement of the burst cue, and as a result we would have to adjust the low ranking of the Cue constraint *[ᵊ]/ /.[8] For lack of experimental evidence, we assumed (based on acoustic descriptions of Italian plosives) that Italian listeners will perceive an auditory input without a final plosive burst, i.e., [ɛ ˺_], as not containing a surface plosive. Future perception experiments will need to support or falsify this assumption, too, and might again lead to a possible adjustment of the native perception grammar: If the experimental data show that our assumption was incorrect and Italian listeners rely far less on the plosive burst than we suggested, then the Cue constraint *[˺_]/k/ will need to be lower ranked, and Italian naive listeners are then predicted to have little problems in perceiving Korean unreleased final plosives as plosives.

[8] We actually expect the perception of a surface schwa by Italians to partly depend on the place of articulation of the plosive, with a very salient velar burst resulting in more schwa perceptions (as in this case the additional schwa-like formants are less important and probably less often occurring) than for less-salient alveolar or bilabial bursts.

As we can see from this, the formalization of a perception grammar is not only informed by experimental data, but it in turn informs experiments, by creating testable hypotheses for future perception experiments.

4.2 Speech perception is not that categorical

The perception grammars we set up in §3 resulted in categorical behaviour, with only one winning candidate per auditory input. In reality, listeners exhibit more variable behaviour, as the non-categorical results of perception experiments show. This variation can be due to several factors, as mentioned already in the discussion of the three languages in §2. Linguistic factors such as preceding vowel quality, stress placement, etc., could be directly implemented with separate Cue constraints that, e.g., differentiate between the mapping of a burst cue onto a surface plosive after tense and after lax vowels. Variation that is not due to such linguistics factors could be within and across speakers (see e.g. Zhou & Hamann 2020 for an illustration of both inter- and intra-speaker variation in L2 perception).

Intra-speaker variation can be dealt with by employing Stochastic OT (Boersma & Hayes 2001): Rather than a ranking order, constraints have values on a ranking scale, where two constraints that are closely ranked are likely to switch positions at evaluation time when noise is added onto the ranking values. This could, e.g., be used to implement the variable perception of coda plosives without burst release in AE. As shown in §3.1, the perception of an auditory input without burst results in a perceived coda consonant if the constraint *[˺_]/ / is high and the constraint *[˺_]/k/ low ranked, as in Korean, while the reverse ranking results in a percept without a consonant, as in Italian. A close ranking of these two constraints within one grammar, where the actual ranking values have to be determined with a computer simulation and actually-observed frequency distributions (see e.g. Boersma & Hayes 2001), could then account for variation between the two forms in AE.

Inter-speaker variation, on the other hand, could be handled by different constraints rankings, i.e. different grammars. A grammar with a high-ranked *[˺_]/k/ could e.g. be employed for AE speakers who show a clear preference for released plosives in coda position, while a grammar where this constraint is low ranked could account for AE speakers who prefer unreleased coda plosives.

4.3 Default interpretation of cues

We hope that our study also drew attention to the problematic nature of the term “illusory vowel” that is used to refer to a perceptual interpretation that assumedly departs from an expected, default interpretation, namely the absence of vowel-like formant structure in the auditory form as an absence of a surface vowel category. As linguists we need to be aware of the fact that there is no default in the interpretation of auditory material, but that this interpretation is strongly influenced by and optimized for the environment we are exposed to and grew up in (as is the case with any sensory input).

There seems to be, however, a cross-linguistic preference to interpret the presence of some cues as separate phonological units, in our example short vowel-like formant structures as full vowels. This idea has been captured for L2 perception as Recoverability Principle by Weinberger (1994; discussed by Jaggers & Baese-Berk 2020: EL512), according to which salient cues are "preserved" as their own phonological entity. In terms of Cue constraints, this would translate into the principle that each (salient) cue that is present in the auditory signal should be mapped onto a corresponding, separate phonological category. We saw that Italian is not in line with this principle when interpreting vowel-like formant structures. The principle would also be violated in cases where one phonological category has several auditory cues, as e.g. vowel transitions, silent closure and burst all cuing a single plosive. These examples illustrate that the Recoverability Principle can be a tendency in (L2) speech perception, at most.

Acknowledgements

We would like to thank the members of the EmpiPhon research group at the ACLC, the audience at the Epenthesis Workshop in Stony Brook (September 2021) and at the Old-World Conference in Phonology in San Sebstian (January 2022), especially Mishko Bozhinoski, Ellen Broselow, Edoardo Cavirani, Piero Cossu, Stuart Davis, Noam Faust, Michaela Watkins and Chao Zhou, for their valuable comments and questions. We would also like to thank two anonymous reviewers for discussion and feedback.

References

Abramson, Arthur S. & Kalaya Tingsabadh. 1999. Thai final stops: Cross-language perception. *Phonetica* 56(3-4). 111–122.

Bafile, Laura. 2002. Il trattamento delle consonanti finali nel fiorentino: Aspetti fonetici. In *Atti delle XXIII Giornate G. F. S. November 28-30*, 205–212. Pisa: ETS.

Bafile, Laura. 2003. Le consonanti finali nel Fiorentino e nel Napoletano. *Rivista Italiana di Dialettologia. Lingue, Dialetti, Società* 26. 149–178.

Bafile, Laura. 2005. Struttura sillabica e consonanti finali in varietà italiana. *Quaderni del Dipartimento di Linguistica - Università di Firenze* 15. 1–25.

Bauer, Robert S. & Paul K. Benedict. 1997. *Modern Cantonese phonology.* Berlin: De Gruyter Mouton. DOI: 10.1515/9783110823707.

Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117(12). 1989–2054. DOI: 10.1016/j.lingua.2006.11.004.

Boersma, Paul. 2009. Cue constraints and their interactions in phonological perception and production. In Paul Boersma & Silke Hamann (eds.), *Phonology in Perception*, 55–110. Berlin: De Gruyter.

Boersma, Paul. 2011. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In Anton Benz & Jason Mattausch (eds.), *Bidirectional Optimality Theory*, 33–72. Amsterdam: Benjamins. DOI: 10.1075/la.180.

Boersma, Paul & Silke Hamann. 2009. Loanword adaptation as first-language phonological perception. In Calabrese Calabrese & Willem Leo Wetzels (eds.), *Loan Phonology*, 11–58. Amsterdam: Benjamins. DOI: 10.1075/cilt.307.02boe.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1). 45–86.

Broniś, Olga. 2016. Italian vowel paragoge in loanword adaptation. Phonological analysis of the Roman variety of Standard Italian. *Italian Journal of Linguistics* 28. 25–68.

Busà, Maria Grazia. 2008. Teaching prosody to Italian learners of English: Working towards a new approach. In C. Taylor (ed.), *Ecolingua: The Role of E-corpora in Translation, Language Learning, and Testing*, 113–126. Trieste: Edizioni Università di Trieste.

Byrd, Dani. 1993. 54,000 American stops. *UCLA Working Papers in Phonetics* 83. 97–116.

Cavirani, Edoardo. 2015. *Modeling phonologization: Vowel reduction and epenthesis in Lunigiana dialects.* Utrecht: LOT.

Chang, Charles B. 2018. Perceptual attention as the locus of transfer to nonnative speech perception. *Journal of Phonetics* 68. 85–102.

Cook, Thomas L. 1969. Efik. In Elizabeth Dunstan (ed.), *Twelve Nigerian Languages.* 35–46. London: Longmans Green.

Crystal, Thomas H. & Arthur S. House. 1988. Segmental durations in connected-speech signals: Current results. *The Journal of the Acoustical Society of America* 83(4). 1553–1573. DOI: 10.1121/1.395911.

Davidson, Lisa. 2011. Characteristics of stop releases in American English spontaneous speech. *Speech Communication* 53(8). 1042–1058. DOI: 10.1016/j.specom.2011.05.010.

Dorman, Michael F., Lawrence J. Raphael & Alvin M. Liberman. 1979. Some experiments on the sound of silence in phonetic perception. *Journal of the Acoustic Society of America* 65(6). 1518–1532. DOI: 10.1121/1.382916.

Dorman, Michael F., Michael Studdert-Kennedy & Lawrence J. Raphael. 1977. Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics* 22(2). 109–122.

Dupoux, Emanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25. 1568–1578. DOI: 10.1037/0096-1523.25.6.1568.

Durvasula, Karthik, Ho-Hsin Huang, Sayako Uehara, Qian Luo & Yen-Hwei Lin. 2018. Phonology modulates the illusory vowels in perceptual illusions: Evidence from Mandarin and English. *Laboratory Phonology* 9. 7. DOI: 10.5334/labphon.57.

Durvasula, Karthik & Jimin Kahng. 2015. Illusory vowels in perceptual epenthesis: The role of phonological alternations. *Phonology* 32. 385–416. DOI: 10.1017/S0952675715000263.

Durvasula, Karthik & Jimin Kahng. 2016. The role of phrasal phonology in speech perception: What perceptual epenthesis shows us. *Journal of Phonetics* 54. 25–34. DOI: 10.1016/j.wocn.2015.08.002.

Escudero, Paola & Paul Boersma. 2003. Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In Sundha Arunachalam, Elsi Kaiser & Alexander Williams (eds.), *Proceedings of the 25th Annual Penn Linguistics Colloquium (Penn Working Papers in Linguistics 8.1)*, 71–85.

Flege, James Emil & James Hillenbrand. 1987. A differential effect of release bursts on the stop voicing judgments of native French and English listeners. *Journal of Phonetics* 15(2). 203–208. DOI: 10.1016/S0095-4470(19)30548-0.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18(1). 7–44.

Fowler, Carol & Jonathan Housum. 1987. Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26. 489–504. DOI: 10.1016/0749-596X(87)90136-7.

Grice, Martine, Michelina Savino & Timo B. Roettger. 2018. Word final schwa is driven by intonation: The case of Bari Italian. *The Journal of the Acoustical Society of America* 143(4). 2474–2486. DOI: 10.1121/1.5030923.

Hale, Mark & Madelyn Kissock. 2007. The phonetics-phonology interface and the acquisition of perseverant underspecification. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, 81–101. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199247455.013.0004.

Hale, Mark & Charles Reiss. 2000. 'substance abuse' and 'dysfunctionalism': Current trends in phonology. *Linguistic Inquiry* 31(1). 157–169. DOI: 10.1162/002438900554334.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Halle, Morris, George W. Hughes & J.-P. A. Radley. 1957. Acoustic properties of stop consonants. *The Journal of the Acoustical Society of America* 29. 101–116. DOI: 10.1121/1.1908634.

Hamann, Silke. 2011. The phonetics-phonology interface. In Nancy Kula, Engbert Doede Botma & Kuniya Nasukawa (eds.), *Continuum Companion to Phonology*, 202–224. London: Continuum.

Householder, F. W. Jr. 1956. Unreleased PTK in American English. In Horace Lunt, Hugh McLean & Cornelis H. van Schooneveld (eds.), *For Roman Jakobson*, 235–244. The Hague: Mouton.

Hutin, Mathilde. 2014. Can release burst be responsible for epenthesis? A case study of Korean. In *Proceedings of the 5th International Conference on Phonology and Morphology*. Chonnam National University, Gwangju.

Iverson, Gregory K. & Ahrong Lee. 2006. Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13. 49–87. DOI: 10.1075/kl.13.03gki.

Jaggers, Zachary & Melissa M. Baese-Berk. 2020. Investigating a bias for cue preservation in loanword adaptation. *JASA* 147(6). EL511–EL516. DOI: 10.1121/10.0001375.

Kabak, Bariş & William J. Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50(1). 23–52. DOI: 10.1177/00238309070500010201.

Kang, Hyunsook. 1996. English loanwords in Korean. *Studies in Phonetics, Phonology and Morphology* 2. 21–47.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20. 219–273.

Kim, Hyunsoon. 1998. A phonetic characterization of release and nonrelease: The case of Korean and English. *Language Research* 34. 347–368.

Kim, Jungyeon. 2022. Multinomial regression modeling of vowel insertion patterns: Adaptation of coda stops from English to Korean. *Linguistics* 60(6). 1989–2017. DOI: 10.1515/ling-2021-0031.

Kim-Renaud, Young-Key. 1987. Fast speech, casual speech and restructuring. *Harvard Studies in Korean Linguistics* 2. 341–359.

Kirby, James. 2011. Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41(3). 381–392. DOI: 10.1017/S0025100311000181.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic grammar – A formal multi-level connectionist theory of linguistic well-formedness: An application. In *12th annual conf. CSS pod*, 884–891. Psychology Press.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation.* Cambridge: MIT Press.

Lisker, Leigh. 1999. Perceiving final voiceless stops without release: Effects of preceding monophthongs versus nonmonophthongs. *Phonetica* 56(1-2). 44–55.

Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20. 384–422. DOI: 10.1017/S0025100311000181.

Lisker, Leigh & Arthur S. Abramson. 1970. The voicing dimension: Some experiments in comparative phonetics. In *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague.

Martin, Samuel. 1951. Korean phonemics. *Language* 27. 519–533.

McQueen, James & Anne Cutler. 1997. Cognitive processes in speech perception. In William J. Hardcastle & John Laver (eds.), *The Handbook of Phonetic Sciences*, 566–586. Oxford: Blackwell.

Miatto, Veronica. 2020. Perception of word-final inserted vowels and syllabicity in Italian. In *Proceedings of the 11th International Conference of Experimental Linguistics (ExLing 2020)*, 137–140.

Miatto, Veronica. 2022. Between intrusivity and epenthesis: A case of Italian word-final inserted vowels. Manuscript, Stony Brook University.

Miatto, Veronica, Silke Hamann & Paul Boersma. 2019. Self-reported L2 input predicts phonetic variation in the adaptation of English final consonants into Italian. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*, 949–953. Melbourne: International Phonetic Association.

Ogasawara, Naomi & Natasha Warner. 2009. Processing missing vowels: Allophonic processing in Japanese. *Language and Cognitive Processes* 24. 376–411. DOI: 10.1080/01690960802084028.

Paradis, Carole & Darlene LaCharité. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33. 379–430.

Parker, Frank & Thomas Walsh. 1981. Voicing cues as a function of the tense/lax distinction of vowels. *Journal of Phonetics* 9. 353–358. DOI: 10.1016/S0095-4470(19)30969-6.

Passino, Diana. 2008. *Aspects of consonantal lengthening in Italian*. Padua: UniPress.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137–157. Amsterdam: Benjamins.

Polivanov, Evgenij Dmitrievič. 1931/1974. La perception des sons d'une langue étrangè. In Evgenij Dmitrievič Polivanov (ed.), *Selected works: Articles on general linguistics*, 223–237. The Hague: Mouton.

Repetti, Lori. 2012. Consonant-final loanwords and epenthetic vowels in Italian. *Catalan Journal of Linguistics* 11. 167–188.

Repp, Bruno H., Alvin M. Liberman, Thomas Eccardt & David Pesetsky. 1978. Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance* 4(4). 621.

Rositzke, Harry. 1943. The articulation of final stops in General American speech. *American Speech* 18. 39–42. DOI: 10.2307/487266.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 219–250. San Diego: Academic Press.

Storto, Luciana & Didier Demolin. 2002. The phonetics and phonology of unreleased stops in Karitiana. *Annual Meeting of the Berkeley Linguistics Society* 28(1). 487–497.

Urua, Eno-Abasi E. 2004. Ibibio. *Journal of the International Phonetic Association* 34(1). 105–109.

Weinberger, Steven. 1994. Functional and phonetic constraints on second language phonology. In Mehmet Yavas (ed.), *First and second language phonology*, 283–302. San Diego: Singular.

Zhou, Chao & Silke Hamann. 2020. Cross-linguistic interaction between phonological categorization and orthography predicts prosodic effects in the acquisition of Portuguese liquids by L1-Mandarin learners. *Proceedings of Interspeech 2020*. 4486–4490.

Zhou, Chao & Silke Hamann. 2024. Modelling the acquisition of the Portuguese tap by L1-Mandarin learners: A BiPhon-HG account for individual differences, syllable-position effects and orthographic influences in L2 speech. *Glossa* 9(1). 1–39. DOI: 10.16995/glossa.9692.

Chapter 10

Prokaryotic syllables and excrescent vowels in two Yuman languages

Martin Krämer
UiT The Arctic University of Norway

Excrescent vowels in two Yuman languages (Cocopa and Jamul Tiipay) and the phonotactic restrictions for their occurrence show that vowels that fulfil some criteria of excrescent vowels are not always a phonetic reflex without repercussions for syllabification (Hall 2006), but rather signal the presence of an additional, albeit non-canonical, syllable. They are inserted in syllables without a nucleus/mora, which renders them inaccessible for higher level prosodic computation. Degenerate, minor or semisyllables, i.e., syllables without a nucleus, have elsewhere been postulated for stray consonants that add a beat/mora accessible to foot construction. The two Yuman languages discussed here add to the typology of minor syllables by contributing minor syllables with an onset and an optional coda, but without a nucleus or a mora. They also provide evidence for a second type of intrusive vowel.

1 Introduction

Long sequences of consonants defying the Sonority Sequencing Generalisation (Selkirk 1984, Zec 1988, 2007, Clements 1990) are attested in a number of languages. They have been analyzed in a range of ways from syllabic, moraic to unsyllabified or in an appendix position (Bagemihl 1991, Lin 1997, Ridouane 2008, Vaux & Wolfe 2009, Zimmermann 2013 among others). Some of the consonant sequences in Cocopa (Yuman; Crawford 1966) are described as interrupted by excrescent vocalic offglides. This matches descriptions of similar sequences in other languages with alleged syllabic obstruents, such as Tashlhiyt Berber or Georgian (see Easterday 2019 for an overview). Crawford (1966) analyzes these

Martin Krämer. 2024. Prokaryotic syllables and excrescent vowels in two Yuman languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 225–245. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264546

as syllables that just contain an onset or an onset and a coda but no phonological nucleus. Such syllables without a nucleus have independently been proposed by McCarthy & Prince (1990), Broselow (1992), Shaw (1994) and Repetti (1994). New arguments for this analysis for Cocopa come from recent insights into the distinction between excrescent and epenthetic vowels. Excrescent vowels are determined in their quality by surrounding consonants, and are shorter and less prominent than epenthetic and lexical, i.e., phonologically present vowels (Hall 2006, 2011; see also the discussion in Easterday 2019).

If Crawford's analysis is adopted, the schwa vowels or vocalic offglides are excrescent vowels without a phonological affiliation. Hall (2006)'s assertion that such intrusive vowels are not accompanied by additional syllable structure cannot be upheld (see also Hall 2024 [this volume], for further refinement of the typology). The excrescent vowels in Cocopa and Jamul Tiipay, which do not have a prosodic association or even a root node, are placed in and signal the presence of an additional syllable – albeit one without a nucleus or mora, a prokaryotic syllable. I present data from Cocopa and its sister language Jamul Tiipay in evidence of the presence of an additional syllable containing the consonant(s) separated from a cluster by an apparently excrescent vowel and conclude that the prosodic invisibility of such syllables must be caused by the lack of a nucleus and mora rather than the lack of a syllable.

Cocopa displays sonority sequencing defying sequences that are not broken up by excrescent vowels, such as presented in (1a), as well as those that receive such a vocalic offglide (1b) and (1c). The choice is determined by the presence of intervening sonorants that then serve as codas of the prokaryotic syllables, in avoidance of syllable-internal sonority roller coaster rides, as in (1c) or other phonological criteria, such as the avoidance of sequences of identical manner.

(1) Cocopa consonant clusters (Crawford 1966; accent indicates stress)[1]

a.	ps̠kʷá	'I gossip about him'
	ksca.ʔárk	'dry!'
	scxuʔá:k	'she hangs up several (things)'
b.	ɬʲpᵃm.wák	'you are to ride him'
	rⁱxúp	'tin can'
c.	pᵃmⁱnṭⁱmá:k	'we abandon them'

[1] Transcriptions from both sources, Crawford (1966) and Miller (2001), are adapted to IPA by the author. The excrescent vowels are transcribed as superscript vowels also indicating their quality as transcribed and described by Crawford in Cocopa examples. In examples from Jamul Tiipay, the excrescent vowels are transcribed as schwa, corresponding to the <e> used in Miller's orthographic transcriptions.

The presence of uninterrupted clusters as well as those broken up by vocalic offglides suggests that Cocopa employs both appendixes/complex onsets as well as degenerate syllables to prosodify its consonant sequences.

The situation is similar in the related language Jamul Tiipay (Miller 2001), even though Miller does not claim that they signal the presence of prokaryotic syllables. Jamul Tiipay displays optional as well as obligatory vowel insertion. Vowel insertion is obligatory if sonorants are involved (2a) and optional between obstruents (2b), suggesting that the sonority requirements on onsets of degenerate syllables are stricter than on full syllables. Sequences of sibilant and stop are never broken up (2c), suggesting that Tiipay also permits appendixes/complex onsets, though in a much more restricted fashion than Cocopa.

(2) Jamul Tiipay consonant clusters

a.	/m-ʃ-ja:j/	məʃəja:j	‘to be afraid’
	/kʷ-n-ma:w/	kʷənəma:w	‘his/her father’s mother’
b.	/x-t̪at̪/	xt̪at̪ /xət̪at̪	‘(someone’s) back’
	/k-ʃ-uː-pit/	kʃu:pit / kəʃu:pit	‘close it!’
	/t̪-t̪-k-juːt̪/	t̪ət̪əkjuːt̪	‘to greet (pl)’
c.	/ʃ-puk/	ʃpuk	‘to lay head on pillow’
	/s-pir/	spir	‘to be strong’
	/s-t̪u/	st̪u	‘to pick up, gather, get’
	/s-kan/	skan	‘to flee’

Tiipay inserted schwa, described as a “non-organic vowel” by Miller (2001), is also not stressable, as is claimed to be typical for excrescent vowels.

The excrescent vowels in Yuman are a phonetic side effect of adjustments in syllabic structure to integrate excess consonants into syllables and they are present in non-canonical or prokaryotic syllables. Unstressability is either an effect of the absence of a nucleus (and accordingly a mora) or just the absence of a mora and thus a unit that can receive stress or can be recognized in mora or syllable counting for foot formation.

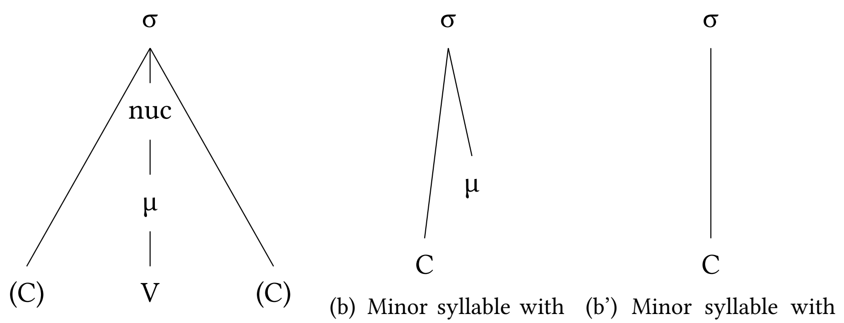
The paper is organized as follows. In Section 2 I provide the theoretical background by first summarizing the state of the art of degenerate syllables, concluding that the more appropriate term is prokaryotic, and then discussing the distinction between excrescent and epenthetic vowels. Section 3 first provides relevant background information on Cocopa and Jamul Tiipay phonology and morphology. It continues with a more detailed discussion of the nature and location of excrescent vowels and argues for strict phonotactics of prokaryotic syllables in Cocopa and more strict phonotactics of prokaryotic syllables in Jamul

Tiipay. Section 4 puts the degenerate syllables proposed for Yuman into a larger typological and theoretical context and Section 5 concludes.

2 Theoretical background

2.1 Degenerate syllables

Degenerate, minor or semisyllables have been proposed in several empirical contexts. McCarthy & Prince (1990) propose to re-analyze superheavy syllables at the right word edge in Arabic as a heavy maximally bimoraic syllable plus an extrametrical syllable consisting of a consonant only. Broselow (1992) adds to this by postulating degenerate syllables with an onset and those with a rhyme. A moraic consonant in a degenerate consonantal syllable is detected by Repetti (1994) at the end of some words in Friulian. Discussing syllable phonotactics and reduplication in Mon Khmer languages, Shaw (1994) proposes to enrich the impoverished model of the syllable in Moraic Theory (Hayes 1989, Zec 1995) by reintroducing the nucleus constituent. In full syllables this constituent is obligatory and associated with at least one mora (Figure 1a). In minor syllables, that is, stray consonants or consonant sequences, the nucleus is missing (Figure 1b) or both nucleus and mora are absent (Figure 1b').



(a) Full syllable

(b) Minor syllable with missing nucleus

(b') Minor syllable with missing nucleus and mora

Figure 1: Full and minor syllables (Shaw 1994)

These defective syllables can be referred to as templates in reduplication processes, as well as the prosodification of excess consonants that do not fit into regular syllables for violating Sonority Sequencing (Clements 1990), or because they show properties such as compensatory lengthening or otherwise unexpected stress placement that justify their analysis as a separate syllable.

Cho & King (2003) propose the term semisyllable, defined as a syllable without a mora and without a coda. They list the following properties of semisyllables.

(3) Properties of semisyllables (Cho & King 2003: 187)
   a. No nucleus
   b. No codas
   c. No stress/accent/tone
   d. Prosodically invisible
   e. Well-formed onset clusters (observing SSP)
   f. Restricted to morpheme peripheral positions

What all these proposals have in common is that the syllable types proposed lack a nucleus, which is why the term *prokaryotic syllable* is more appropriate than the familiar terms *degenerate, minor* or *semisyllable*. To my knowledge, the term prokaryotic has not been proposed before. Given that the consonants in such a syllable are subject to different phonotactic restrictions, as are the onset and the coda of major or full syllables, i.e., those that have a nucleus, it can be assumed that some more subsyllabic structure is present. This goes beyond what Shaw so carefully proposed and is also more than what Cho & King propose, since, as I will show, prokaryotic syllables may have a coda. Cho & King listed the absence of codas as one of the characteristics of minor syllables. Accordingly, the constituents rime and coda are given in parentheses in Figure 2 to indicate their optionality.

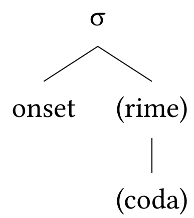


Figure 2: Prokaryotic syllable

2.2 Epenthetic and excrescent vowels

Hall (2024 [this volume]) observes that predictable or inserted vowels behave differently in different contexts and languages. Furthermore, some are invisible to phonological processes such as stress placement or tone assignment. The differences line up into properties that are typical of what she calls intrusive vowels on the one side, and epenthetic vowels on the other. Epenthetic vowels are phonological in the sense that they are affiliated with phonological structure, such as a mora or a syllable nucleus, while excrescent vowels are not. She gives the following catalogue of properties.

(4) Properties of phonologically invisible inserted vowels (intrusive vowels) (Hall 2006: 391)

a. The vowel's quality is either schwa, a copy of a nearby vowel or influenced by the surrounding consonants.
b. If the vowel copies the quality of another vowel over an intervening consonant, that consonant is a sonorant or guttural.
c. The vowel generally occurs in heterorganic clusters.
d. The vowel is likely to be optional, have a highly variable duration or disappear in fast speech rates.
e. The vowel does not seem to have the function of repairing illicit structures. The consonant clusters in which the vowel occurs may be less marked, in terms of sonority sequencing, than clusters which surface without vowel insertion in the same language.

According to Hall, intrusive vowels are not stressable/do not affect stress placement and do not participate in other phonological processes. This is why they are assumed to be devoid of syllable structure. She contrasts these properties with those she attributes to properly phonological epenthetic vowels:

(5) Properties of phonologically visible inserted vowels (epenthetic vowels) (Hall 2006: 391)

a. The vowel's quality may be fixed or copied from a neighboring vowel. A fixed-quality epenthetic vowel does not have to be schwa.
b. If the vowel's quality is copied, there are no restrictions as to which consonants may be copied over.
c. The vowel's presence is not dependent on speech rate.
d. The vowel repairs a structure that is marked, in the sense of being cross-linguistically rare. The same structure is also likely to be avoided by means of other processes within the same language.

Furthermore, phonologically visible inserted vowels occupy a syllable nucleus, usually are moraic and participate in stress assignment and other phonological processes (e.g., vowel harmony).

I will show in the following sections that the excrescent vowels of Cocopa and Jamul Tiipay have most of the properties of intrusive vowels, including phonological invisibility, but do signal the separate syllabification of a consonant sequence that would be an illicit tautosyllabic cluster. Consonants are separated by an intrusive vowel when they cannot coinhabit an onset. The separate syllables created for these consonants are not stressable, are ignored in reduplication, but are subject to constraints of syllable phonotactics. We are thus dealing with prokaryotic syllables of the type outlined in Section 2.1 that consist of an onset and optionally a coda but no nucleus or mora.

3 Excrescent and epenthetic vowels in Yuman

The two languages Cocopa and Jamul Tiipay both belong to the Yuman family. Cocopa is spoken by around 400 people who live north and south of the borders between Mexico, California and Arizona. Jamul Tiipay or just Tiipay is spoken in a neighboring area west of the Cocopa area. For 2007, Ethnologue reported approximately 100 remaining speakers. Diegueño or Kumeyaay, which will be discussed briefly at the end of Section 3.1, is spoken north of Tiipay north and south of the border between California and Mexico. In the 1990s there were an estimated 50 native speakers. All Cocopa and Jamul Tiipay data used here come from Crawford (1966) and Miller (2001), respectively. The inventories of contrastive segments of the two languages are extremely similar, which is why I will present them together in section 3.1. In this section I also discuss relevant aspects of syllable phonotactics, stress and affixation as well as reduplication. Section 3.2 presents the details of vowel intrusion in Cocopa and their analysis. Section 3.3 presents the vowel intrusion patterns of Jamul Tiipay. Section 3.4 summarizes the section.

3.1 Background

To understand the role of excrescent vowels in these two Yuman languages it is essential to first learn about the basic facts of their phonology and morphology. I will first discuss their segment inventories, with special focus on the phonetics of the vowels, consonant cluster phonotactics and stress and close this subsection with a short discussion of reduplication.

Table 1: Cocopa consonants (adapted from Crawford 1966: 25)

	Labial	Dental	Alveol.	Alveo-palatal	Velar	Labio-velar	Uvular	Labio-uvular	Glottal
Stops	p	t̪	t	c	k	kʷ	q	qʷ	ʔ
Nasals	m	n̪		ɲ					
Affric.				t͡ʃ					
Fricat.		s̪	s	ʃ	x	xʷ			
Lat. fric.			ɬ	ɬʲ					
Lateral			l	ʎ					
Rhotic		r̪	r						
Glides				j		w			

Table 1 displays the Cocopa consonants as described by Crawford.

Cocopa has a slightly larger consonant system than Jamul Tiipay, which lacks the uvular stops and does not distinguish between two rhotics. Crawford also describes an additional coronal stop and fricative for Cocopa.

Table 2 is adapted from Miller (2001: 39ff). Neither language has a laryngeal contrast, but nevertheless both have a sizeable consonant inventory.

Table 2: Jamul Tiipay consonants (adapted from Miller 2001: 39)

	Labial	Dental	Alveol.	Alveo-palatal	Velar	Labio-velar	Glottal
Stops	p	t̪			k	kʷ	ʔ
Nasals	m	n̪		ɲ			
Affricate				t͡ʃ			
Fricatives		s̪		ʃ	x	xʷ	
Lat. fric.			ɬ	ɬʲ			
Lateral			l	ʎ			
Rhotic			r				
Glides				j		w	

Crawford describes three contrastive vowels for Cocopa and Miller discusses the status of a fourth one. Some schwas, she claims, are not predictable and therefore have to be analyzed as present in the lexicon. Table 3 is adapted from Miller (2001: 12).

Crawford (1966: 13) discusses two additional vowels. /e/ occurs in Spanish loans and is consistently mid to upper-mid unrounded, unless it is replaced by

Table 3: Cocopa and Jamul Tiipay vowels (Crawford 1966, Miller 2001)

Front	Central	Back
i, iː		u, uː
	ə	
	a, aː	

Cocopa /i/, which many speakers do. The other vowel, /o/, is only found in one interjection that expresses "frustration or disappointment".

In both languages, the realization of the three contrastive vowels depends considerably on their environment. In Cocopa, the front high vowel is a bit centralized when preceded or followed by an alveolar consonant. The long front vowel is a bit higher than the short one. This difference is not reported for the back high vowels. These vary in height and are lowered slightly when followed by another vowel with only one intervening consonant or when followed by a sibilant. Crawford only gives examples with following low vowels for the first condition. There might actually be some kind of height harmony operative here. The low vowel is lowest when preceding the stressed vowel, only separated from it by one consonant, and slightly raised and fronted when preceded by a palatal consonant and even more so when surrounded by palatals. Elsewhere, it is a central low vowel with the long one a bit lower than the short.

The intrusive vowel is described as a vowel similar to /u/ when followed by a labiovelar consonant, including /w/, an /i/-like vowel when followed or preceded by a palatal or dental before any consonant except the labiovelars, and as a schwa-like vowel in all other environments (Crawford 1966: 38, see also Mansfield et al. 2024 [this volume], for similar environmental colouring of inserted vowels). He transcribes them as superscript *i*, *u* and *a*, respectively.

Miller (2001: 20) gives a similar description of what she calls "inorganic" schwa, the vowel that "is inserted between consonants to break up clusters" in Tiipay: it is never long and never stressed. Its quality is determined by surrounding consonants, resulting in [ɪ], [ʊ], [ə]. Whenever two conditions overlap schwa may be realized as any of the available options, e.g., [ɬʲəxʷiːw] 'skunk'; this vowel could be realized as [ɪ] or [ə] because it is preceded by a palatal consonant (a palatal voiceless lateral fricative) or it can be realized as [ʊ] because it is followed by a labialized velar [xʷ]. If schwa is only separated from the following vowel by a glottal stop, it may be realized as a copy of the short allophone of that vowel. The three lexical vowels vary according to environment in a similar way as in

Cocopa. Since some schwas are unpredictable, such as those in the left column in (6), Miller assumes them to be present in lexical representations. The forms on the right serve to illustrate the same phonotactic context without a schwa. The last form shows a context for optional vowel intrusion, i.e., a prokaryotic syllable.

(6) Unpredictable schwas in Jamul Tiipay (Miller 2001: 21)

aləmi	'beard'	xalma	'gourd rattle'
wanəpu	'buttocks'	xənpaɬ	'tongue'
xəmuk	'to be three'	xəmuɬ / xmuɬ	'to be foamy'

Cocopa complex onsets can consist of almost any sequence of up to four obstruents, with the following restrictions. If a voiceless lateral is involved, it is cluster-initial. If a glottal stop is involved, it is final in the cluster. Sequences of stops are not attested, but sequences of fricatives are. If there is a nasal it is final too. The only examples for complex onsets with a sonorant other than a nasal are loanwords and have the sonorant as the internal member. Complex onsets of unstressed syllables are slightly more restricted. For example, they do not contain a glottal stop.

(7) Cocopa stressed complex onsets

a. Sequences of obstruents

xps̪íw	'be blue, green'
ps̪kʷá	'I gossip about him'
ɬʲksís	unidentified plant species
psc ʔá:w	'I have them as daughters'
scxʔú:ɲ	'yellowshafted flicker'
xs̪á:m	'be almost'
s̪xʈú	'I spit'

b. Rising sonority

ʃmá	'I sleep'
ɬʲmár	'I light a fire'
ɬʲs̪míx	'I intend to lay something big in'
ɬʲjú:m	'I think'
tré:n	'train'
krú:s̪	'cross'

Jamul uninterrupted initial clusters, as already indicated in the introduction, are much more restricted. They all start in a sibilant and the internal consonant is always a stop.

(8) Jamul Tiipay uninterrupted initial consonant clusters

/ʃ-puk/	ʃpuk	‘to lay head on pillow’
/s-pir/	spir	‘to be strong’
/s-tu/	stu	‘to pick up, gather, get’
/s-kan/	skan	‘to flee’
/ʃ-ʈu/	ʃʈu	‘to shove (with hands or instrument)’

All other sequences are broken up at least optionally by a schwa vowel, as will be discussed in the next subsection.

On stress, Crawford (1966: 28) states that there are three levels, strongly stressed, stressed and unstressed. Within what he calls a “macrosegment”, which I interpret to roughly coincide with a word, there is usually only one stressed syllable. However, he also states that within a macrosegment with more than one syllable preceding the stressed syllable, “the first unstressed syllable has a slightly stronger stress than the following unstressed syllables” (Crawford 1966: 29). Crawford is very clear about the unstressability of prokaryotic syllables: “A stressed or strongly stressed syllable can only be one which contains a vowel.” And on syllables he states that “[a] syllable can be entirely consonantal and consist of an onset only or of an onset and a coda with a predictable ‘murmur’ vowel following the onset as phonetic peak.” (Crawford 1966: 34) According to Miller (2001)’s description of stress in Jamul, and in Yuman in general (Langdon 1975), stress on schwa syllables is not an option because of the morphological nature of stress placement. Stress is always placed on the morphological root of a word, which predominantly has the shape (C)V(C). Since the morphology is mostly prefixing, this results in word-final stress in most cases. In the few cases of bigger roots, as illustrated in (8), stress is still on the last vowel and schwas are not stressed.

Cocopa verb root reduplication is a semi-productive process. I consider it semi-productive because many reduplicated verbs do not have a non-reduplicated base form. Many do, however, and we can observe some regularities that indicate that it is impossible to reduplicate consonantal syllables. The preferred verb root for reduplication is of the form CVC(C). Roots with complex codas can be reduplicated, while verbs with complex onsets are not reduplicated. Initial consonant clusters arise only when an instrumental prefix is added to the reduplicated form. In this case inflection for person is possible, while the other reduplicated forms are impersonal uninflected forms. Inflection is realized on an adjacent auxiliary. One instrumental prefix exemplified is of the form CV- and the other is a sibilant. This sibilant causes either deletion of the reduplicant-initial (root) consonant or its alternation from a fricative into a stop, creating either a simple onset or a cluster, adhering to the restrictions for such clusters found in Jamul (8).

(9) Cocopa prefix-reduplicant interaction
ʃírmír 'I take aim' (probably from *mírmír i* 'to be straight')
skárxár/sxárxár 'I break into small pieces'

Thus, none of the many consonantal prefixes that would create complex onsets or prokaryotic syllables are reduplicated or even used to further derive or inflect reduplicated forms.

Jamul reduplication is similarly restricted and unproductive. The base is maximally CVCC, as in (10a) and (10b). Of the 23 reduplicated verb stems Miller collected, only two forms have a prefix, and even there it is prefixed to the reduplicated form, as shown in (10c).

(10) Reduplication in Jamul Tiipay
a. milmil 'to be narrow'
b. aʃkaʃk-i 'to go up and down, back and forth'
c. t͡ʃəxəlxul 'to gargle' cf. təkəlkul 'to pile (things) up'

Verbs such as [txiːl] 'to get dressed, wear clothes' do not seem to undergo reduplication, not even with a reduced reduplicant (e.g., *[xiːl-txiːl])

The alternations observed in Cocopa (9) suggest that the restriction of reduplication on verb roots with simple onsets is a phonological one, and that the many stems that are formed with derivational consonantal prefixes, such as causatives, do not undergo reduplication because complex onsets and minor syllables are banned in the reduplicant.

Compare these reduplication patterns with those found in the sister language Diegueño/Kumeyaay (Langdon 1966), which displays almost identical patterns of schwa insertion. Langdon reports to have found many reduplicated forms. Almost all reduplicate only the stem syllable, as in the other two languages. There is, however, a very small set of bisyllabic reduplicants. Interestingly, in three of the four forms Langdon found, the initial vowel is a schwa.

(11) Bisyllabic reduplicants in Kumeyaay (Langdon 1966: 202)
kułaːɬ kułaːɬ 'to go up and down (like when riding a horse)'
ɬəxup ɬəxuːp 'holes all over' (cf. ɬəxup 'hole, cave')
səkap səkaːp 'half and half, to be more than half full' (cf. səkap 'to be half')
xəkaɬ xəkaːɬ 'to be scalloped, uneven at the edges, to have teeth missing'

The fact that there are bisyllabic reduplicants in Kumeyaay and that they contain schwas can be counted as weak evidence that the schwas in Kumeyaay are phonological, unlike those in Cocopa and Jamul Tiipay. Presumably, reduplication targets prosodic structure, such as moras.

As illustrated in (7) and (8), the restrictions on initial clusters differ in Cocopa and Jamul. In the following we will more closely examine the insertion sites for intrusive vowels in both varieties and conclude that also the phonotactic constraints on prokaryotic syllables differ slightly.

3.2 Cocopa intrusive vowel landing sites

The intrusive vowel prevents a sonority rise and consecutive fall, as illustrated in (12). Minor syllables in Cocopa can have complex onsets, just like major syllables, i.e., containing two obstruents (12a). A vowel is inserted if either a sonorant is followed by an obstruent or vice versa (12b). The forms in (12b) could theoretically be syllabified with fewer inserted vowels, i.e., fewer prokaryotic syllables, as indicated by the conceivable but unattested forms marked with a question mark in (12b). This would, however, result in obstruent codas followed by sonorant onsets, as the question marked forms show. Such rising sonority profiles across syllables violate the Syllable Contact Law (Murray & Vennemann 1983), according to which sonority should fall from one syllable to the next. Clusters with variable intrusion sites involve sonorants and either syllable contact created is wellformed (12c).

(12) Prokaryotic syllable phonotactics

a.	sx^{a}m.pá	‘yellowjacket’	
	ɬ^{j}p^{a}m.wák	‘you are to ride him’	
	pʃk^{u}.wáːk^{x}	‘we intend to return him’	
b.	m^{a}.k^{i}.ɲáːp	’you relate’	?m^{a}k.ɲáːp
	p^{a}.m^{i}n.ʈ^{i}.máːk	’we abandon them’	?p^{a}m.n^{i}ʈ.má:
	m^{i}.c^{i}m.p^{a}.káːwc	’you meet each other’	?m^{i}c.m^{i}p.káːwc
c.	ɲ^{i}m.ɲ^{i}.k^{w}ájs / ɲ^{i}.m^{i}ɲ.k^{w}ájs	‘we are your mother’s brothers’	
	ɲ^{i}ɬ^{j}.mwa.jáːc / ɲ^{i}.ɬ^{ji}m.wa.jáːc	‘you are around in it’	

As in major syllables, sequences of stops are avoided (unless the last stop is a glottal stop). Obstruent-sonorant sequences are avoided too, suggesting that rising sonority in complex onsets is marked and restricted to loanwords.

(13) Stop-stop and obstruent-sonorant onsets are avoided

ʈ^{i}.ʈʔá:p	‘I turn something upside down’
p^{a}.q^{i}.la.ʃáw	‘he cleaned him’
ʈ^{i}.má:j	‘waves of the ocean’
p^{i}.lík	‘I taste’

Cocopa allows both complex onsets as well as, presumably, appendix plus onset initial clusters, with the appendix filled with a sibilant. Clusters that exceed these structures with maximally three consonants, and clusters that do not conform to the sonority restrictions and the manner OCP banning consecutive stops, are divided up into prokaryotic syllables, which can have complex onsets and codas. The minor syllables are more restricted in Jamul Tiipay as are regular syllables.

3.3 Jamul Tiipay prokaryotic syllable phonotactics

The schwa vowel emerges between stops and between sonorants, between stops and sonorants but not glides, between sonorants and obstruents, and sonorants and sonorants (14a). It does not occur between sibilants and stops (8), but between sibilants and glides, as illustrated in the second example in (14). (14b) shows that also string-internally sibilant-stop sequences are tolerated, as word-initially (8).

(14) Jamul Tiipay cluster resolution

a.	/t-ɲur/	təɲur	‘to curl (hair), to decorate’
	/m-ʃ-ja:j/	məʃəja:j	‘to be afraid’
	/k^{w}-n-ma:w/	k^{w}ənəma:w	‘his/her father’s mother’
	/t-t-k-ju:t/	tətəkju:t	‘to greet (pl)’
	/ɲ-ʃ-k-ʔ-mak/	ɲəʃkəʔmak	‘s/he took it away from me’
b.	/k-s-kan/	kəskan	‘run away!’
	/m-ʃ-t-u:-jaj/	məʃtu:jaj	‘to be afraid (pl)’
	/ɲ-ʃ-p-a:-ʔ-ʔá:w-a/	ɲəʃpaʔá:wa	‘they made us stand up’

The glottal stop behaves differently in that it can precede any consonant but not follow a consonant.

(15) No insertion between glottal stop and other Cs

/ɲkʔ-wi:w/	ɲəkəʔwi:w	‘look at me!’
/ɲ-ʃ-k-ʔ-mak/	ɲəʃkəʔmak	‘s/he took it away from me’

Unlike Crawford, Miller distinguishes between obligatory schwa insertion (14 and 15) and optional intrusion (16). Miller does not always give two forms in these examples. From the fact that she lists them in this context, I conclude that they display optional schwa.

(16) Variable epenthesis between two voiceless obstruents

/x-tat/	xtat /xətat	‘(someone’s) back’
/p-ʔaw/	pʔaw	‘to stand, step; (for rain) to fall’
/t-k-aː-xaːp/	tkaːxaːp	‘bracelet’
/k-ʃ-uː-pit/	kʃuːpit/ kəʃuːpit	‘close it!’
/t͡ʃxlkaj/	t͡ʃxəlkaj	‘kidneys’
/t͡ʃ-k-uːjaw-a/	t͡ʃkuːjawa	‘to teach’
/s-naːj/	snaːj	‘to dip up (water)’
/t͡ʃ-mi/	t͡ʃmi	‘to lay (long or large object) down’

In longer sequences we find fewer intrusive vowels than expected, suggesting that there are limits to the number of prokaryotic syllables in a row. The occasional transcription of an optional schwa in string-internal sequences of sonorants followed by obstruents shows that internal clusters are preferably syllabified tautosyllabically as onsets and not heterosyllabically as coda-onset sequences. However, it is also noteworthy that none of the given forms has a word-initial complex onset to a prokaryotic syllable. Word-initial clusters are all followed by a full vowel. This differs from Cocopa, where we find up to three consonants followed by an intrusive vowel (12a).

(17) Bigger clusters

/t-t-xʷak/	tətxʷak	‘to break (brittle object) (pl)’
/m-ɲ-kurʔak/	məɲkurʔak	‘your husband’
/m-ʎ-piʃ/	məʎpiʃ/məʎəpiʃ	‘you are small’
/t͡ʃ-k-piːk/	t͡ʃəkpiːk	‘to squash many’
/t-k-xap/	təkxap	‘to put on, wear (bracelet, ring, shirt, eyeglasses)’
/k-t-k-xap/	kətkəxap/kətəkxap	‘put (bracelet, ring, shirt, eyeglasses) on!’
/m-m-ʃ-jaːj/	məmʃəjaːj məməʃəjaːj	‘you are afraid’
/ɲk-m-ʃ-ʔ-jaːj/	ɲəkəmʃəʔjaːj	‘be afraid of me!’
/m-m-ʃ-kʷaɬʲ/	məmʃəkʷaɬʲ/məməʃkʷaɬʲ	‘you bother him/her; s/he bothers you’
/ɲm-m-ʃ-kʷaɬʲ/	ɲəməməʃkʷaɬʲ	‘you bother me’

Miller (2001) analyzes the cluster patterns with an across-the-board schwa epenthesis rule and several schwa deletion rules that apply optionally, resulting in optionality for most of the inserted schwas. In addition, these rules also apply either rightwards or leftwards in a cluster. Thus, the two forms of 'you bother him/her; s/he bothers you' in (17) are the result of rightward or leftward application of a late schwa deletion rule scanning the result of the general schwa insertion rule |mə-mə-ʃə-k^{w}aɬ^{j}|. We can thus add another property of intrusive invisible vowels, optionality (4d).

3.4 Summary

The stress and reduplication patterns of Cocopa and Jamul Tiipay do not show any sign that the syllables with intrusive vowels are accessible by higher level prosody. Closer inspection of the distribution of intrusive vowels and where they are optional and where not reveals restrictions on onsets and codas, which differ between regular and prokaryotic syllables as well as across the two languages. Cocopa regular onsets display complexity, allowing for proper complex onsets with rising or plateauing sonority and an OCP constraint that bans sequences of stops as well as sequences of fricatives. Prokaryotic syllables in Cocopa can have complex onsets consisting of obstruents alternating in continuancy and they can have simple codas. Regular onset phonotactics are stricter in Jamul Tiipay, permitting basically only sibilant-stop clusters and prokaryotic syllables seem to allow complex onsets only under duress, i.e., to avoid codas or too many consecutive prokaryotic syllables.

4 Theoretical implications

The Yuman patterns discussed here contribute to the understanding of minor syllables on the one hand and of excrescent vowels on the other.

4.1 Non-canonical syllables

I will not try to integrate all proposals of consonants dominated only by a syllable node or a mora or no syllable structure or an appendix into one model of a typology of prokaryotic syllables. These are in many cases competing proposals. However, a few remarks are in order.

The appendix (e.g., Vaux & Wolfe 2009 and references cited there) is an additional position that can be attached to the syllable, preceding the onset or a higher-level prosodic category, such as the foot or the word. In the analysis of

Indo-European languages, this is usually invoked to account for sibilant-initial clusters that violate sonority sequencing. Clusters of rising sonority are treated as complex onsets. The only segment allowed in this position is accordingly a sibilant, and usually only one, as in English. The first segment in complex onsets is usually an obstruent and the second segment is a sonorant or in more restrictive languages a non-nasal sonorant. We have seen the former, i.e., sibilant-initial clusters, in Jamul Tiipay. They are not only highly restrictive in the first position, allowing only sibilants, but also in the second position, in which we find only stops. The cross-linguistically widespread rising sonority clusters are not attested or only in recent loanwords. In Cocopa, we find more combinations of obstruents in initial clusters. Only consecutive stops are avoided. Such obstruent clusters can have more than two members.

An analysis of the sibilant-initial clusters in Jamul Tiipay as appendix plus simple onset is an obvious choice. Whether the Cocopa obstruent sequences are appendixes plus onset or complex onsets with a strict requirement for flat and low sonority is a more intricate issue. We will not solve this here, since these elaborations only serve to rule out an appendix analysis for the clusters broken up by excrescent vowels. If an intrusive vowel can or must be inserted between an appendix and the following onset consonant, it should also be attested in the sibilant + stop sequences in Jamul Tiipay. This is not the case. The vowel intrusion patterns can thus not be analyzed by the stipulation of an appendix position for each consonant preceding an intrusive vowel. Assuming that several appendixes can precede the first syllable of a word, one would also not expect any clusters of two consonants inside a sequence of appendixes, as in the second form in (12b) or those in (12c). This suggests that there is more elaborate syllable structure than just a sequence of appendixes.

The Yuman prokaryotic syllable is also different from the minor syllables that have been proposed as the weak part of the sesquisyllable in Southeast Asian languages (Matisoff 1973, Shaw 1994, for a recent discussion see Butler 2014). The minor (or half) syllable in a sesquisyllabic word precedes a full or major syllable. These arguably form iambic feet together and the minor syllable might even bear a tone (Svantesson & Karlsson 2004, Butler 2014). The minor syllable in sesquisyllables thus has to have prosodic structure that makes it visible for footing and that makes it a licit tone-bearing unit. The mora is usually assumed to be the relevant unit in iambic feet and tone association.

Yuman prokaryotic syllables contribute neither to foot construction nor can they be said to be prosodically active in any other way. Their only purpose is to avoid illicit consonant sequences within canonical syllables.

4.2 Excrescence and epenthesis

As has been argued at length here, the schwa vowels of Cocopa and Jamul Tiipay emerge between the constituents of syllables without nuclei. They are not inserted to repair an illicit structure, but they do emerge as a side effect of such an adjustment in avoidance of marked or ungrammatical phonotactics. They thus do not comply with one of the central criteria invoked by Hall for the diagnosis of excrescent/intrusive vowels. They do, however, fulfil other important criteria. They are short, reduced, unstable, and variable, and their quality is dependent on that of neighboring consonants, i.e., their immediate phonetic neighborhood.

Considering how the schwas of Cocopa and Jamul Tiipay have properties of both phonetic and phonological inserted vowels and lack properties of each type, it is tempting to propose the existence of a third category. Using the existing terminology, we could distinguish between three types of inserted vowels – epenthetic vowels that are integrated in the phonological structure, excrescent vowels that signal additional phonological structure, and intrusive vowels, which are phonologically irrelevant. The latter two categories are both phonologically invisible in that they are not available for the phonology, e.g., to house tones, contribute to foot construction or participate reliably in vowel harmony. However, this would be premature. The most important distinction is whether a vowel is affiliated with phonological structure, i.e., parsed in a syllable nucleus, or not. In addition, I think one can reasonably argue that the emergence of the excrescent vowels is a side effect of gestural timing, similar to that of Hall's intrusive vowels.

The presence of the additional onset to house the first consonant(s) of a sequence that would not constitute a well-formed onset with this initial consonant, has phonetic consequences. The final consonant in an onset is expected to have a substantial release phase, especially if it is a stop, which is expected to turn into a vowel, which always occupies the nucleus of a major or canonical syllable. The consonant in a prokaryotic syllable is also in an onset, it is just not followed by a nucleus. One can, however, assume that the articulatory targets are determined by its position in the syllable structure and that it behaves in the same way as a consonant in a major syllable does. The only difference is that the onset consonant in a prokaryotic syllable is not followed by a nucleus and thus no phonological vowel. It is thus the automatic articulatory mechanics at the end of the consonant in an onset that makes observers perceive a vocalic offglide or schwa-like vowel. A coda consonant on the other hand is not expected to have much of a release. In many languages, word-final stops do not have an audible release at all. The articulatory "habits" for onsets thus facilitate the emergence of a following transitional or excrescent vocoid, while the articulation patterns for coda consonants do not easily provide space for such a non-phonological vocoid.

An excrescent vowel in a prokaryotic syllable is thus not the result of an overlap of the transition between two consonants and the opening gesture of a phonological vowel. Its perception is, however, the result of gesture coordination determined by the prosodic structure associated with the surrounding segments and thus a phonetic by-effect of abstract phonological representations.

Accordingly, it is more appropriate to divide inserted vowels into epenthetic and excrescent vowels, and the latter into those caused by mere gestural coordination and those caused by the mapping of abstract phonological representations to articulatory actions.

5 Conclusions

Excrescent vowels in Cocopa emerge in response to sonority fluctuation inside consonant clusters and thus signal the presence of a prokaryotic syllable. Jamul Tiipay schwa insertion and variability is similarly conditioned by the sonority of surrounding consonants. Hall's main argument for assuming the consonants flanking excrescent vowels to not project a separate syllable was their inactivity in stress placement and other prosodic patterns. This inactivity of prokaryotic syllables is explained here by their prosodic deficiency causing their inability to contribute to higher level prosodic structure: Prokaryotic syllables, i.e., those with an optional excrescent vowel, do not have a nucleus and do not project a mora, as proposed by Cho & King. In contrast to what Cho & King stipulated, however, Cocopa prokaryotic syllables can have a coda. If the distinction between obligatory and optional schwa observed by Miller in Jamul Tiipay signals a phonological difference, we are most probably dealing with defective syllables which contain a nucleus but do not project a mora in the case of obligatory schwas. Thus, they too are ignored in the computation of stress, feet, or other syllable counting operations.

There are two main results of this study. First, with the help of excrescent or intrusive vowels, we can recognize prokaryotic syllables. These syllables are inaccessible for prosodic computation because they lack a nucleus and a mora, but they are syllables because they are subject to syllable phonotactic constraints on onsets and codas. Cho & King's definition of what they call semisyllables thus must be broadened to include prokaryotic syllables with a coda.

Second, there are two types of intrusive (non-phonological) vowels. The first type are those described by Hall as stemming from gestural overlap of the vowel within a syllable with the transitions between consonants within that syllable. The second type, described here, emerges as a phonetic effect of standard articulatory patterns in the realization of consonants in specific syllable positions, in this case the rightmost consonantal position in an onset.

References

Bagemihl, Bruce. 1991. Syllable structure in Bella Coola. *Linguistic Inquiry* 22. 589–646.

Broselow, Ellen. 1992. Parametric variation in Arabic dialect phonology. In Ellen Broselow, Mushira Eid & John McCarthy (eds.), *Perspectives on Arabic linguistics IV*, 7–45. Amsterdam: John Benjamins.

Butler, Becky. 2014. *Deconstructing the Southeast Asian sesquisyllable: A gestural account*. Cornell University. (Doctoral dissertation).

Cho, Young-mee Yu & Tracy Holloway King. 2003. Semisyllables and universal syllabification. In Caroline Féry & Ruben van de Vijver (eds.), *The syllable in Optimality Theory*, 183–213. Cambridge: Cambridge University Press.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In John Kingston & Mary Beckman (eds.), *Between the grammar and physics of speech* (Papers in Laboratory Phonology 1), 283–333. New York: Cambridge University Press.

Crawford, James Mack. 1966. *The Cocopa language*. University of California at Berkeley. (Doctoral dissertation).

Easterday, Shelece. 2019. *Highly complex syllable structure: A typological and diachronic study* (Studies in Laboratory Phonology 9). Berlin: Language Science Press. DOI: 10.5281/zenodo.3268721.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429. DOI: 10.1017/S0952675706000996.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2). 253–306.

Langdon, Margaret. 1966. *A grammar of Diegueño: The Mesa Grande dialect*. University of California at Berkeley. (Doctoral dissertation).

Langdon, Margaret. 1975. Boundaries and lenition in Yuman languages. *International Journal of American Linguistics* 41. 218–233.

Lin, Yen-Hwei. 1997. Syllabic and moraic structures in Piro. *Phonology* 14. 403–436.

Mansfield, John, Rosey Billington & Hywel Stoakes. 2024. Vowel predictability and omission in Anindilyakwa. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 57–84. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264532.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.), *Consonant types and tone (Southern California Occasional Papers in Linguistics No. 1)*, 73–95. Los Angeles: University of Southern California.

McCarthy, John & Alan Prince. 1990. Prosodic morphology and templatic morphology. In Mushira Eid & John McCarthy (eds.), *Perspectives on Arabic linguistics II (Papers from the Second Annual Symposium on Arabic Linguistics)*, 1–54. Amsterdam: John Benjamins.

Miller, Amy. 2001. *A grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter.

Murray, Robert & Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59(3). 514–528.

Repetti, Lori. 1994. Degenerate syllables in Friulian. *Linguistic Inquiry* 25(1). 186–193.

Ridouane, Rachid. 2008. Syllables without vowels: Phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25. 321–359.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: The MIT Press.

Shaw, Patricia. 1994. The prosodic constituency in minor syllables. In Erin Duncan, Donka Farkas & Philip Spaelti (eds.), *Proceedings of the 12th West Coast Conference on Formal Linguistics*, 117–132.

Svantesson, Jan-Olof & Anastasia Karlsson. 2004. Minor syllable tones in Kammu. In Bernard Bel & Isabelle Marlien (eds.), *Proceedings, International symposium on tonal aspects of languages (TAL2004)*, 177–180. Chinese Academy of Sciences.

Vaux, Bert & Andrew Wolfe. 2009. The appendix. In Eric Raimy & Charles E. Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 101–143. Cambridge: MIT Press.

Zec, Draga. 1988. *Sonority constraints on prosodic structure*. Stanford University. (Doctoral dissertation).

Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12. 85–129.

Zec, Draga. 2007. The syllable. In Paul V. De Lacy (ed.), *The Cambridge handbook of phonology*, 161–195. Cambridge: Cambridge University Press.

Zimmermann, Eva. 2013. Vowel deletion as mora usurpation: The case of Yine. *Phonology* 30. 125–163.

Chapter 11

On the diachrony of lateral epenthesis

Michael Ramsammy
University of Edinburgh

This paper discusses the phenomenon of lateral epenthesis from a diachronic perspective. Within the theoretical context of the life cycle of phonological processes, I briefly discuss two cases that are frequently mentioned in general studies on epenthesis, namely English /l/-intrusion and /l/-epenthesis in Motu. Shifting the focus to Indo-Aryan, /l/-insertion in Hindi causative verbs is then discussed in detail. After tracing diachronic developments through Old and Middle Indo-Aryan, I argue that Hindi /l/-causatives are not the product of a historical pattern of allophony in which [p] surfaces after vowel-final bases in causative forms. Similarly, I show that evidence for an analogical development of /l/-causatives, as has been assumed within the philological tradition, is relatively weak. Consequently, I propose an alternative account under which Hindi /l/-causatives emerge due to sonority-driven optimisation of historical /j/-epenthesis.

1 Introduction

Theoretical accounts of epenthetic patterns involving consonants aim to respond to two main questions. The first is that of what constrains epenthesis: i.e. what environmental and structural factors trigger the insertion of consonantal material, and how is this determined on a language-specific basis. The second question is what consonants can be epenthetic: i.e. what specific qualities or feature values favour the occurrence of particular consonants in epenthesis environments. Research around this second theme has revealed typologically common patterns, and also some that are much less well attested.

This paper discusses a relatively rare phonological phenomenon, namely lateral epenthesis. Being rare, cases of lateral epenthesis are typically mentioned

Michael Ramsammy. 2024. On the diachrony of lateral epenthesis. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 247–270. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264548

in more general theoretical discussions of epenthesis, which may be either language-specific or cross-linguistic in nature. In this paper, the focus is given to the diachronic emergence of epenthetic laterals. I first discuss some examples that are better known in the epenthesis literature, specifically cases in English and Oceanic languages (Section 1). Thereafter, the core of the paper is dedicated to Indo-Aryan. In Section 2, I present data from Hindi causative verbs in which /l/ occurs preceding the first causative suffix, /-ɑ/, after vowel-final bases: e.g. /so-/ 'sleep' vs [sʊlɑ] 'put to sleep'. As a hiatus-resolution strategy, /l/-insertion presents explanatory challenges in both synchronic and diachronic dimensions.

Section 3 of the paper traces the morpho-phonological history of causative formations in Indo-Aryan and discusses proposals that have previously been put forward to explain the occurrence of Hindi /l/-causatives. These rely strongly on the assumption of analogical processes of change. Whilst these claims are philologically well-founded, it is also generally acknowledged that the emergence of lateral epenthesis in causative verbs remains somewhat mysterious: as Bloch (1965: 141) puts it, "the real history of these suffixes is hidden from us".

In view of this, Section 4 presents a reconstruction of a pathway of change leading to the development of lateral epenthesis over time. The analysis is situated theoretically within the life cycle of phonological processes (Bermúdez-Otero 2015, Bermúdez-Otero & Trousdale 2012, Iosad 2020, Ramsammy 2015, Roberts 2012, Sen 2016, Turton 2016, 2017). Thus, the overarching aim is to construct a plausible scenario through which the development of lateral epenthesis in Indo-Aryan can be explained on the basis of core phonological principles and without exclusive reliance on analogy.

1.1 Cases of lateral epenthesis

1.1.1 English

Gick (1999) describes the behaviour of intrusive /l/ in American English in the context of /ɔ/ and /a/ (see also Gick 2002). In varieties such as those spoken by working-class communities around Philadelphia, /l/-intrusion operates in a similar manner to intrusive /r/, which has been more widely discussed in the phonological literature (e.g. Barras 2011, Hall 2013, Mompeán-González & Mompeán-Gillamón 2009, Sóskuthy 2013, Uffman 2007, inter alia). Gick highlights the relationship between /l/-vocalisation and /l/-intrusion.[1] Where /l/ undergoes vo-

[1]Relatedly, Johnson & Britain (2007) highlight the relationship between the development of vocalised variants of /l/ in British dialects and pre-existing clear~dark lateral allophony. There are, however, cases of /l/-vocalisation in other languages that do not depend on an allophonic pattern of this type: e.g. Cibaeño Spanish (Alba 1979).

calisation in coda position, e.g. [drɔː] *drawl*, but is retained intervocalically, e.g. [drɔːlɪŋ] *drawling*, intrusive /l/s emerge subsequently in non-etymological contexts: e.g. [drɔːlɪŋ] *drawing*, [braːlɪz] *bra is*.[2]

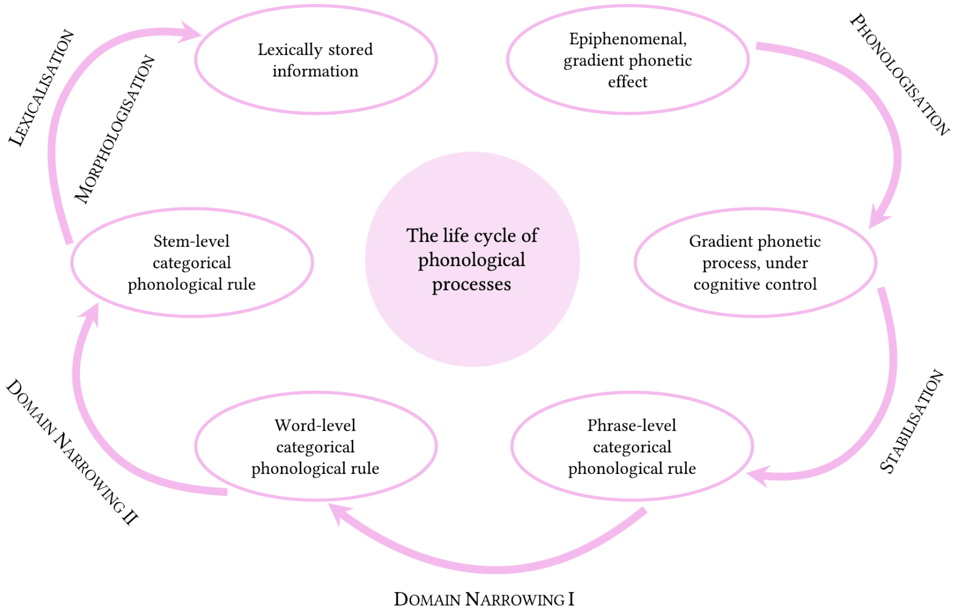


Figure 1: The life cycle of phonological processes (adapted from Ramsammy 2015: 38)

/l/-intrusion of this type represents a classic case of rule inversion (Vennemann 1972, Bermúdez-Otero & Börjars 2016) that is also consistent with the core assumptions of the life cycle of phonological processes. The life cycle is illustrated in Figure 1: the initial stages are of particular relevance here. In brief, a fundamental claim of the life cycle is that all language-internal phonological innovations begin as low-level phonetic phenomena that are below the level of speaker awareness. Over time, this can change: a gradient phonetic effect can transcend the level of speaker awareness, and thus, come under the speaker's

[2]Some speakers do not epenthesise /l/ in examples like *bra is* (cf. Bermúdez-Otero & Börjars 2016: 741). As Gick (2002: 178) notes, this pattern is observed for speakers who merge /ɔ/ and /ɑ/, hence [dɹɑːlɪŋ]. Further to this, Weissmann (1970) describes a similar pattern of /l/-linking in Bristol English, e.g. [æfrɪkələneɪʃə] *Africa and Asia*, which is an example of a broader /l/-*epithesis* pattern (Gick 1999) that causes homophony between pairs such as *idea~ideal*, *Eva~evil*, etc. (Trudgill 1999).

cognitive control. This is the process of *phonologisation*. Following phonologisation, the innovation is still gradient in nature. In a further phase of development known as *stabilisation*, it can evolve into a discrete operation, governed by e.g. featural or representational changes. The latter phases of the life cycle involve changes to the domain of application of the phonological innovation, which gradually shrink over time through *domain narrowing*. The final phase is *lexicalisation*, which involves restructuring of underlying representations, or alternatively *morphologisation*, which involves the conversion of a phonological operation into a morphological one.

In this connection, the development of /l/-epenthesis is illustrated in Table 1. Before the onset of change, /l/ occurs freely both in word-medial and word-final positions. At Stage 1, /l/ in word-final position begins to weaken: Gick notes that this involves partial vocalisation through undershoot or a reduction in magnitude of the tongue tip gesture. As predicted by the life cycle, /l/-weakening is a gradient phenomenon in this phase. It then proceeds through the phonologisation and stabilisation phases leading to Stage 2, at which point /l/-vocalisation is reinterpreted as a discrete phonological rule.

Table 1: English lateral epenthesis by rule inversion

Stage 0:	
Consonantal coda /l/	Consonantal onset /l/
/drɔːl/ ⟶ [drɔːɫ]	/drɔːl+ɪŋ/ ⟶ [drɔːlɪŋ]
Stage 1:	
Weakening of coda /l/	No word-medial weakening
/drɔːl/ ⟶ [drɔː^{ɫ}]	/drɔːl+ɪŋ/ ⟶ [drɔːlɪŋ]
Stage 2:	
Vocalisation of coda /l/	Preservation of word-medial /l/
/drɔːl/ ⟶ [drɔː]	/drɔːl+ɪŋ/ ⟶ [drɔːlɪŋ]
Stage 3:	
Input restructuring	Reinterpretation: /l/ is epenthetic
/drɔː/ ⟶ [drɔː] (*drawl*)	/drɔː+ɪŋ/ ⟶ [drɔːlɪŋ] (*drawling*)
Stage 4:	
Underlying ∅	Extension of /l/-epenthesis
/drɔː/ ⟶ [drɔː] (*draw*)	/drɔː+ɪŋ/ ⟶ [drɔːlɪŋ] (*drawing*)

The critical reanalysis happens at Stage 3, which can be assumed to be a learner-driven development. A child acquiring the grammar at this stage reaches the generalisation that, rather than an underlying /l/ vocalising word-finally, the underlying form for *drawl* in fact contains no /l/. The child therefore postulates that the [l] in forms like *drawling* is epenthetic. The Stage 2 deletion rule has thus undergone inversion to a Stage 3 epenthesis rule. The final phase in the development is the generalisation of this rule to all VV sequences (whether derived through affixation, as in the case of *drawing*, or dialect-dependently, through concatenation of words into phrases, as in *the bra is*).

1.1.2 Oceanic

In addition to English, Vaux (2001: 7) and Morley (2015: 15) mention Motu (Oceanic, Papua New Guinea), which displays instances of word-initial /l/ that are present neither in cognate forms in related languages nor in Proto-Oceanic reconstructed forms. Data (adapted from Blust 1999) are given in Table 2. As shown, elimination of word-initial onsetless syllables through epenthesis occurs in both Motu and Fijian in the context of the vowel /a/. Blust comments that Motu /l/ is a reflex of historical /j/, such that the Fijian data here can be said to resemble a prior phonological stage for Motu.

Table 2: /l/-epenthesis in Motu and /j/-epenthesis in Fijian

	Proto-Oceanic	Motu	
a.	ansan	lada	‘name’
b.	asaŋ	lada	‘gills’
c.	apaʁat	lahara	‘North-West wind and season’
d.	api	lahi	‘fire’
e.	aja	lala	‘father’s sister; woman’s brother’s child’
f.	aku	lau	‘I’
	Proto-Oceanic	**Fijian**	
g.	ansan	jaca	‘name’
h.	asaq	jaca	‘grate, grind, sharpen’
i.	aɲam	janajana	‘loosely plaited’
j.	ane	jane	PO: ‘termite’; Fij. ’moth species’
k.	asi	jasi	‘sandalwood’

Unlike /l/-epenthesis in English, Oceanic /j/-epenthesis is not the result of rule inversion. Instead, we might reconstruct the change as the result of vowel breaking. Under this scenario – and in keeping with the life cycle's fundamental claim that phonological changes begin as low-level phonetic effects – the inception of the change might have involved onsetless word-initial syllables being articulated with something like a short onglide: e.g. [ᵊasi]. As this pattern then underwent phonologisation and rose above the level of speaker awareness, it could gradually have become more /j/-like: i.e. involving a larger articulatory movement and a greater acoustic distance from the original /a/.[3] Following the life-cycle trajectory, a rule of /j/-insertion before /a/ would have emerged in the grammar as the gradient phonetic pattern stabilised historically: i.e. /asi/ ⟶ [asi] > [ᵊasi] > [ʲasi] > [jasi].

The additional innovation in Motu is an extension of this pathway of change, i.e. ∅ > /j/ > /l/. What, then, could cause lateralisation of the historical epenthetic /j/? This development is most probably perceptually driven. Whereas phonologisation and stabilisation of vowel breaking may lead to the emergence of word-initial glides, glides make relatively poor onsets because of their acoustic similarity to vowels.[4] This fact underlies the prohibition that some languages enforce on syllable-initial glides. It is also the cause of fortition effects, such as glide hardening, which reduce the sonority of glides typically in syllable-initial positions (cf. (4) below). Within the OT literature, this has been formalised as a markedness hierarchy for onset sonority (Gopal 2018, Gouskova 2004).[5]

(1) Onset sonority markedness hierarchy (adapted from Gopal 2018: 84)
*Ons/j ≫ *Ons/r ≫ *Ons/l ≫ *Ons/n ≫ *Ons/z ≫ *Ons/d ≫ *Ons/s ≫ *Ons/t

As indicated, high-sonority glides (represented in (1) by the shorthand /j/) are highly marked when they occur in syllable-initial position, whereas low-sonority voiceless stops (represented here by /t/) are optimal onsets. In this regard, (1) is relevant for Fijian too. Blust (1999: 10–11) argues that /j/-epenthesis in Fijian was a lexically gradual change. The words listed in Table 2 are relatively late examples. Words that were early targets for epenthesis show evidence of subsequent

[3]A parallel to this development is the Great Vowel Shift, specifically the emergence of /aɪ/ and /aʊ/ from historical /iː/ and /uː/, respectively. This is thought to have involved vowel breaking and progressive stages of change in which the vocalic onset became more phonetically distinct from the original high vowel over time: e.g. [iː] > [iɪ] > [əɪ] > [aɪ] (cf. Krug 2017: 246ff.).

[4]Note that the etymological /j/ in Table 2e also lateralises in Motu. In addition to /lala/, Blust lists the Motu forms /huala/ 'crocodile' and /mala/ 'tongue', from Proto-Oceanic /puqaja/ and /maja/, respectively.

[5]See Bokhari (2024 [this volume]) on the implementation of a similar hierarchy in the analysis of Hijazi Arabic vowel epenthesis.

glide hardening, i.e. /j/ > /c/. In the examples in Table 3, the historical epenthetic /j/ was maximally optimised with regard to onset sonority. In accordance with (1), hardening of /j/ to /c/ maximally increases the perceptual distinctiveness between the onset and following vowel.

Table 3: Epenthesis and glide hardening in Fijian

	Proto-Oceanic	Fijian	
a.	asam	caca	‘fern species’
b.	aŋin	cagi	‘wind’
c.	aŋo	cago	PO ‘yellow’; Fij. ‘turmeric’
d.	aʁu	cau	‘shore tree’
e.	apaʁat	cava	‘North-West monsoon, storm wind’

In Motu, optimisation of historical epenthetic /j/ was not as extreme. Rather than reducing the sonority of the epenthetic onset to the lowest possible level, as in Fijian, the change /j/ to /l/ in Motu represents a sonority decrease of just two increments in the onset sonority hierarchy. This confirms the fact that the threshold of onset well-formedness with regard to sonority is set language-specifically in regard to (1). Indeed, I shall argue that sonority reduction from /j/ to /l/ is of relevance for Hindi too: this is discussed in Section 4. Before proceeding to discussion of that phenomenon, Section 2 below first introduces the Hindi data that form the basis of the remainder of the paper.

2 Hindi causative verbs

Modern Standard Hindi forms causative verbs with two suffixes, /-ɑ/ and /-ʋɑ/. /-ɑ/ forms the so-called first or direct causatives, and /-ʋɑ/ is used for the second or indirect causatives. Some verbs select only one suffix or the other: for example, /putʃʰ/ ‘ask’ can take the second formative (i.e. /pʊtʃʰ-ʋɑ/ ‘cause to be asked’) but not the first (*/pʊtʃʰ-ɑ/). A second subset of verbs can take both suffixes, thereby yielding semantically distinct triplets (cf. Bhatt & Embick 2017): e.g. /kud/ ‘jump’ ~ /kʊd-ɑ/ ‘bounce’ ~ /kʊd-ʋɑ/ ‘make jump’.

In this paper, I focus exclusively on the first causative forms. As illustrated in the data below, formation of first causatives triggers a range of phonological processes. Table 4 lists verb stems containing the lax vowels, /ɪ, ʊ, ʌ/, and /ɛ, ɔ/, which are derived from historical diphthongs. In these verbs, the stem vowel does not alternate following affixation of causative /-ɑ/. However, stems containing the tense vowels, /i, u, ɑ/, or /e, o/ – which, unlike reference varieties of English, pattern as lax vowels in Hindi – are subject to neutralisation in causative derivations. As shown in Table 5, stems containing tense vowels display laxing in their corresponding causative forms, whereas /e/ and /o/ raise to [ɪ] and [ʊ], respectively, as in the examples in Table 6.

Table 4: Non-alternating stems in /ɪ, ʊ, ʌ, ɛ, ɔ/

	Stem		Causative stem	
a.	lɪkʰ	‘write’	lɪkʰ-ɑ	‘dictate’
b.	sʊn	‘hear’	sʊn-ɑ	‘tell, cause to hear’
c.	ʧʌl	‘move, go’	ʧʌl-ɑ	‘drive’
d.	pʰɛl	‘stretch’	pʰɛl-ɑ	‘extend, cause to stretch out’
e.	dɔɽ	‘run’	dɔɽ-ɑ	‘urge on, cause to run’

Table 5: Laxing: stems in /i, u, ɑ/

	Stem		Causative stem	
a.	ʧʰin	‘snatch’	ʧʰɪn-ɑ	‘cause to snatch’
b.	gʱum	‘turn, tour’	gʱʊm-ɑ	‘tell, cause to hear’
c.	mɑn	‘accept’	mʌn-ɑ	‘persuade’

Table 6: Vowel raising: stems in /e, o/

	Stem		Causative stem	
a.	leʈ	‘lie down’	lɪʈ-ɑ	‘make lie down’
b.	bol	‘speak’	bʊl-ɑ	‘call, invite’

Note that the stems listed in Tables 4–6 are all consonant-final. By contrast, the verbs listed in Table 7 have stem-final vowels. Their causative forms in Table 7a–e are of principal interest because /l/ occurs as a hiatus-breaker between the base and suffix vowels. Whilst the laxing patterns are not critical for the operation of /l/-insertion, vowel neutralisation can also be observed here: /i/ laxes to [ɪ] in examples Table 7a–b and the stems containing the mid vowels /e/ and /o/ display raising in their causative counterparts in Table 7c–e.

Table 7: /l/ after vowel-final bases in first causative forms

	Stem		Causative stem	
a.	pi	‘drink’	pɪ-lɑ	‘water, irrigate’
b.	si	‘sew’	sɪ-lɑ	‘caused to sew’
c.	de	‘give’	dɪ-lɑ	‘cause to be given’
d.	ro	‘cry’	rʊ-lɑ	‘cause to cry’
e.	so	‘sleep’	sʊ-lɑ	‘put down to sleep’

What is particularly interesting about these forms is that /l/-insertion is not a regular hiatus-breaking strategy in Hindi. The forms in Tables 8 and 9 provide a point of comparison. In Table 8, hiatuses formed by the addition of future-tense suffixes to vowel-final bases are tolerated on the surface. Ohala (1983: 72–74) notes that /j/-epenthesis in these types of VV-sequences is sometimes observed (e.g. [pɪjegɑ], [sɪjegɑ], etc.), but this is not obligatory.[6]

Table 8: Hiatus after vowel-final bases in future forms. Formation of future forms triggers laxing of stem /i/ in Table 8a–b in the same way that causative /-ɑ/ does in Table 5a and Table 7a–b. The same pattern is noted with stem /u/ in Table 8g. /de/ ‘give’ is irregular, but note that hiatus after stem-final /ɪ/ is unrepaired, as shown.

	Stem		2PL.FUT.INFOR	3SG.FUT.M
a.	pi	‘drink’	pɪ-og-e	pɪ-eg-ɑ
b.	si	‘sew’	sɪ-og-e	sɪ-eg-ɑ
c.	de	‘give’	didʒɪ-og-e	didʒɪ-eg-ɑ
d.	ro	‘cry’	ro-og-e	ro-eg-ɑ
e.	so	‘sleep’	so-og-e	so-eg-ɑ
f.	ɑ	‘come’	ɑ-og-e	ɑ-eg-ɑ
g.	ʧʰu	‘touch’	ʧʰʊ-og-e	ʧʰʊ-eg-ɑ

Similarly, hiatuses in stem-suffix sequences occur in a subset of the perfective forms listed in Table 9. In the feminine singular and masculine plural, formed

[6] Glottal stops or [w] may also occur variably in /e+e/ sequences: e.g. [kʰee], [kʰeʔe], [kʰewe] ‘row’ (2SG.FUT.SUBJ). Like variable /j/-insertion, these patterns may be partially dependent on dialectal factors that remain to be fully explored.

through suffixation of /-i/ and /-e/, respectively, hiatus is observed after vowel-final stems. Interestingly, this is not the case in the masculine singular forms. Here, addition of perfective /-ɑ/, which is syncretic with causative /-ɑ/, triggers the emergence of a pre-suffixal [j]. Note that [j] is obligatory in these forms (unlike optional /j/-epenthesis in forms like [pɪegɑ~pɪjegɑ]) and that [j] never occurs after consonant-final bases, as in Table 9d–e. These facts confirm that it is specifically hiatus before /-ɑ/ that is targeted for obligatory repair in Hindi. In causatives, the outcome is a pre-suffixal [l], whereas in masculine perfectives, it is a pre-suffixal [j].[7]

Table 9: Perfective forms

	Stem		3sg.perf.f	3pl.perf.m	3sg.perf.m
a.	pi	‘drink’	pi	pi-e	pɪ-jɑ, *pi-ɑ
b.	ro	‘cry’	ro-i	ro-e	ro-jɑ, *ro-ɑ
c.	ɑ	‘come’	ɑ-i	ɑ-e	ɑ-jɑ, *ɑ-ɑ
d.	leʈ	‘lie down’	leʈ-i	leʈ-e	leʈ-ɑ, *leʈ-jɑ
e.	bol	‘speak’	bol-i	bol-e	bol-ɑ, *bol-jɑ

It is therefore clear that regular /l/-insertion and /j/-insertion are morphologically conditioned (cf. Vaux 2001). Regarding the causative pattern, I refer to a derivational sketch in Section 4 which assumes that /-lɑ/ is an allomorph of /-ɑ/ in the present-day language. In this sense, and depending upon how restrictively one wishes to define the term “epenthesis”, /-lɑ/-*selection* may be the most appropriate label for the synchronic Hindi patterns outlined in Table 7. Nevertheless, I also shall make the claim that the occurrence of [l] in causative forms is governed by phonological factors. As will also be discussed in Section 4, synchronic morphologically-conditioned allomorphy may reflect patterns of epenthesis that were phonologically active in earlier forms of a language. In this connection, I now turn to discussing some of the key historical facts that are relevant for the Hindi patterns, beginning with causative formations in Sanskrit.

3 Historical considerations

Causative verbs were formed with the suffix /-ɑjɑ/ in Classical Sanskrit, as illustrated in Table 10a–d below (data adapted from Mayrhofer 1978: §129). Vowel-final roots show special behaviour: note that in Table 10e–g, the consonant /p/

[7]Thus, causative perfectives exhibit a double repair: e.g. [pɪlɑjɑ] ‘irrigated’ (3sg.caus.perf.m).

intervenes between the base and suffix vowels. The parallel between pre-suffixal /p/ in these forms and the Hindi causatives with pre-suffixal /l/ listed in Table 7 is clear. However, beyond the fact that they serve a hiatus-breaking function, the phonological correspondence between these consonants is not obvious.

Table 10: Sanskrit causative formation

	Root		Causative forms (3SG.PRES.INDIC)	
a.	kṛ	'do'	kār-ɑjɑ-ti	'causes to do'
b.	ɟɑn-	'be born'	ɟɑn-ɑjɑ-ti	'begets, procreates'
c.	dṛʃ	'see'	dɑrʃ-ɑjɑ-ti	'causes to see'
d.	budh-	'wake'	bodh-ɑjɑ-ti	'causes to wake'
e.	gā	'sing'	gā-pɑjɑ-ti	'causes to sing'
f.	ɟɲā	'know'	ɟɲā-pɑjɑ-ti	'causes to know'
g.	dā	'give'	dā-pɑjɑ-ti	'causes to give'

Regarding the origin of the Hindi pattern, a direct phonological change, /p/ > /l/, seems unlikely for two reasons. Firstly, lateralisation of labial consonants is not attested in Indo-Aryan. Cases of lateral excrescence after labial consonants have been documented in other language families. Most notably, Common Slavic /pj/-clusters developed into /pl^{j}/: this is represented dialectally in examples like Belorussian *kanópli*, Slovenian *konóplja* vs Upper Sorbian *konopje*, Polish *konop~konopie* 'hemp' (see Shevelov 1964: 219ff. for further examples). Baltic shows a similar development, particularly Latvian: e.g. Latvian *pl'aūt* 'mow' vs Lithuanian *piáuti* 'cut' (Endzelin 1922: §84).[8] This raises the question of whether a lateral consonant could have emerged in the /p/ causative allomorph in the history of Indo-Aryan: e.g. /-pɑjɑ/ > $^{?}$/-pjɑ/ > $^{?}$/pl(j)ɑ/ > /-lɑ/. Indeed /pj/- and /pl/-clusters in Hindi are rare (at least outside of English loanwords) relative to the very high frequency of /pr/. However, whereas /pj/ and /mj/ do occur medially in Sanskrit (Masica 1991: 161), these clusters never lateralise in any Indo-Aryan variety. A development resembling Balto-Slavic /pj/ > /pl/ with subsequent loss of /p/ is therefore an unlikely source of the Hindi /l/-causatives.

This connects to the second reason for rejecting /p/ > /l/ as a plausible pathway of change, namely that the /-pɑjɑ/ allomorph was repurposed as the second causative suffix via a development /-pɑjɑ/ > /-pe/ > /-ʋɑ/. This occurred without the emergence of a prevocalic lateral consonant at any recorded stage.

[8]I am grateful to Florian Wandl for bringing these facts to my attention.

In fact, the change /-pɑjɑ/ > /-ʋɑ/ involved intermediate stages that played out differently across Indo-Aryan dialects. For example, /ɑj/-monophthongisation had a significant morpho-phonological impact in the transition from Old (OIA) to Middle (MIA) Indo-Aryan. In Pāli (early MIA), /p/-causatives were retained despite simplification of OIA /-ɑjɑ/ to /-e/: e.g. /ɲā-pē-ti/ 'causes to know', cf. Table 10f – additional examples in Oberlies (2001: §52). By the middle MIA period, this suffix had passed through a phase of intervocalic lenition to /-ʋe/ (e.g. Māhārāṣṭrī /ʤāɳāʋēi/ 'causes to know'), and a lower suffix vowel is observed in later MIA varieties such as Apabhraṁśa: e.g. /kɑrāʋɑi/ 'causes to do' (cf. Bubenik 2003: 230). These changes are also reflected in the patterning of causative suffixes across present-day New Indo-Aryan (NIA) languages, with some exhibiting both first and second causatives that resemble Hindi /-ɑ/ and /-ʋɑ/, and others showing only one type or the other (cf. Masica 1991: §9.6 for a complete overview).[9]

Given the low probability that /l/-causatives originate from historical /p/-causatives, other potential sources have been suggested. One claim is that the Sanskrit form *pālayati* 'saves, brings across' may be the source of Hindi /l/-causatives. *pālayati* is a causative formation on the root /pr̥-/ 'cross over', which, as Jamison (1983: 85) notes, is a variant of the well-attested regular formation, *pārayati*. Dialectal variation between /r/ and /l/ is also well documented for OIA: Norman (2012: 280) comments that "[…] there must have been OIA dialects which turned all -r- and -l- sounds into -l-, others which turned them all into -r-, and still others which mingled the two sounds in different proportions".[10] Thus, despite its status as a dialectal variant of *pārayati*, the argument goes that the pre-suffixal /l/ from *pālayati* generalised as an epenthetic segment that was inserted after vowel-final bases in causative formations. Sen (1960: §156) gives the Apabhraṁśa 1sg causative form *dālayami* (from /dā/, 'give') as a possible instance of this development.

Nevertheless, beyond a very small set of examples, there is no strong evidence from Old or Middle Indo-Aryan that supports an analogical development of /l/-causatives based on a dialectal variant of *pārayati* with /r/-lateralisation. A further point in this connection is that merger of /l/ and /r/ in favour of /l/ is more typical of eastern varieties of Indo-Aryan, for example Eastern Aśokan (early MIA), with Western Aśokan maintaining the /l/~/r/ contrast (Oberlies 2003: 176). This is at least potentially problematic since Modern Hindi evolved from varieties

[9] For example, Odiya has only first causatives in /-ɑ/ and Sinhalese has only /-ʋɑ/. Some varieties, such as Maithili, have second-causative suffixes with /b/ instead of /ʋ/, which presumably arose from voicing but not spirantisation of MIA /-pe/.

[10] Interestingly, /l/ in Kwa (Sande 2024 [this volume]) has multiple possible surface realisations, including [l, ɾ, n].

of MIA, particularly Śaurasenī, which originate in the central-western region.[11] This fact presents a problem also for a second possible historical source of /l/-causatives. A number of Eastern NIA languages exhibit /l/ in perfective forms: examples for 'seen' (from Bloch 1965: 267) include Bengali *dekkhila*, Maithili *dekhal* and Marathi and Oriya *dekhilā*, inter alia. Bloch (1965) identifies this /l/ as a potential source of /l/-causatives. However, /l/ is not observed in perfective forms elsewhere (cf. Hindi /dek^{h}ɑ/ 'seen'). The situation is therefore one of geographical discontinuity: two possible sources of /l/ that could have provided an analogical basis for the development of /l/-causatives in central-western NIA languages are restricted to eastern varieties, both historically and synchronically.

4 An alternative pathway: Epenthesis and onset sonority

4.1 Reconstructing the pre-history of lateral epenthesis

Since evidence for an analogical development is not strongly supported either by historical sources or by the geographical/dialectal facts, it is worthwhile to consider other mechanisms of change that could have led to the Hindi pattern. In this section, I explore an alternative scenario that draws both on the core principles of the life cycle and the analyses of lateral epenthesis in English and Motu already discussed.

Figure 2 illustrates the pathway of change leading to the development of pre-suffixal laterals in Hindi causative verb forms that I shall assume. Stages 1–3 refer to the pre-history of Hindi. As discussed in Section 3, monophthongisation of OIA /-ɑjɑ/ to MIA /-e/ – i.e. the change from Stage 1 to Stage 2 – is well documented. Similarly, textual evidence supports the reconstruction of a later development in which the quality of the MIA causative suffix, /-e/, lowered to /-ɑ/. The post-vocalic allomorph, /-pe/, was similarly affected, in addition to a change in consonantal quality through spirantisation and voicing: i.e. /-pe/ > /-ʋɑ/.

It is then at Stage 4 that the most important innovations for the development of the Hindi pattern happen. Firstly, the /-ʋɑ/ causative suffix that developed from OIA /-pɑjɑ/ acquires a grammatically distinct function: at this point, it is used to form second causatives that are semantically distinct from first causative forms. This would have involved a fundamental reanalysis of /-ʋɑ/, i.e. a *change in status* (Anderson 1981). Whereas /-ʋɑ/ occurs only as an allomorph of the general causative suffix, /-ɑ/, at Stage 3, it has acquired a new status as an independent,

[11]Note that other Western NIA languages such as Panjabi also display /l/-causatives that closely resemble the Hindi forms.

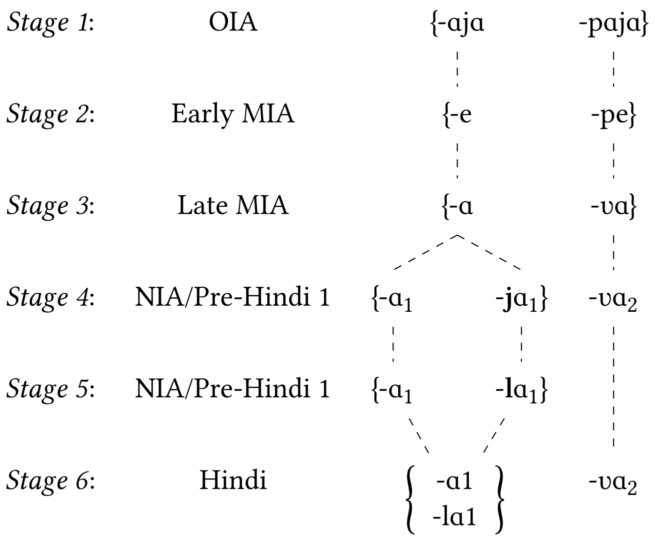


Figure 2: Diachronic development of Hindi causative suffixes

lexically listed verbal suffix by the completion of Stage 4. As noted in Section 3, this development is particular to certain NIA languages; and whilst it is semantic and grammatical in nature, this change entails a crucial phonological consequence, namely that a consonant-initial suffix allomorph is no longer available for first-causative formations.

At this point, it is reasonable to assume that a new phonological repair to VV-sequences generated by affixation of first-causative /-ɑ/ to vowel-final bases would have evolved. In accordance with the reconstruction in Figure 2, one possibility is that this repair could have taken the form of an innovative pattern of /j/-epenthesis in the first instance. This is a somewhat tentative proposal, though it is not without foundation. Similar to the reconstruction of Fijian /j/-epenthesis, such a development may have come about through vowel breaking, or through a gradient liaison of the type /ɑ-ɑ/ ⟶ [ɑ^{j}ɑ] that, as already noted, is commonly observed synchronically in Hindi. Furthermore, verbal stems terminating in /-ɑ/ have a particularly high frequency of occurrence in the language. The life cycle postulates that any phonological innovation begins as a low-level phonetic effect: thus, productions resembling [ɑ^{j}ɑ] may have occurred, for example, as some sort of articulatory or perceptual artifact of speakers producing /ɑ-ɑ/-sequences generated through the formation of causatives.

Whereas the phonetic precursor of the pattern cannot be precisely pinpointed, phonologisation and stabilisation of a gradient pattern of /j/-liaison of this type would lead to a new, categorical rule of /j/-epenthesis specifically targeting hiatus in the context of a suffixal /-ɑ/. Interestingly, under the assumption that a /j/-epenthesis pattern is a plausible precursor to lateral epenthesis, then the developments between Stages 1 and 4 in Figure 2 could be said to represent a kind of "long-distance" rule inversion. In the English patterns discussed in Section 1.1.1, it was observed that lateral epenthesis emerged from a historical pattern of coda /l/-reduction and deletion. Under the scenario presented in Figure 2, loss of /j/ through diachronic truncation of OIA /-ɑjɑ/ to MIA /-e/ is undone at Stage 4: i.e. /j/ is re-supplied through epenthesis in NIA in the same context that it disappeared from in the development from OIA to MIA.

4.2 From /j/-epenthesis to /l/-insertion

The second phase in the development of /l/-insertion occurs at Stage 5. Here, the /j/-epenthesis process innovated at Stage 4 evolves further into a pattern of lateral epenthesis. This parallels what Blust (1999) reconstructs for Motu. In line with the generalisation that /j/ is a poor onset from a perceptual point of view, fortition to /l/ reduces the sonority of the epenthetic material.[12] In the context of the onset sonority hierarchy in (1), this development can be modelled straightforwardly as a diachronic re-ranking of constraints penalising sonorant onsets. This is illustrated in Figure 3.[13]

At Stage 4, a superordinate constraint targeting forms with hiatus enforces a repair by epenthesis to VV-sequences generated by suffixation of causative /-ɑ/ to vowel-final bases like /so/ 'sleep'. The quality of the epenthetic consonant is regulated by lower-ranked constraints. As shown, a constraint such as Dep-seg$_{[+cons]}$ prevents the insertion of [+cons] segments: this eliminates candidate Figure 3a-iii with lateral epenthesis (and any candidate exhibiting additional consonantal material). Accordingly, candidate Figure 3a-ii with /j/-epenthesis is selected as the winner despite its violation of low-ranked *Ons/j. The change from Stage 4 to Stage 5 happens through promotion of *Ons/j. As this constraint militates against syllable-initial /j/, Figure 3b-ii is eliminated by the later grammar. Under this ranking, therefore, the form with lateral epenthesis, i.e. candidate Figure 3b-iii, wins.

In fact, diachronic promotion of a constraint like *Ons/j is well motivated for Indo-Aryan on independent grounds. This is confirmed by a number of other phonological developments that conspired to eliminate /j/-onsets historically. For

[12]The proposal here is therefore for a development that follows from similar principles to Uffmann's (2007) treatment of /r/-insertion in English as intervocalic sonority optimisation.

[13]Figure 3 abstracts away from the vowel-neutralisation process that causes raising of /o/ to [ʊ] in causative formations. I assume that this is governed by other constraints that are omitted from the tableaux here.

/so+ɑ/	*hiatus	Dep-seg$_{[+cons]}$	*Ons/j	*Ons/l
i. [sʊɑ]	*!			
☞ ii. [sʊjɑ]			*	
iii. [sʊlɑ]		*!		*

(a) Stage 4

/so+ɑ/	*hiatus	*Ons/j	Dep-seg$_{[+cons]}$	*Ons/l
i. [sʊɑ]	*!			
ii. [sʊjɑ]		*!		
☞ iii. [sʊlɑ]			*	*

(b) Stage 5

Figure 3: /j/-epenthesis > /l/-epenthesis by re-ranking of onset sonority constraints

example, OIA /j/ hardened in MIA in a similar way to the Fijian change, /j/ > /c/, resulting in a sonority reduction in word-initial contexts. The outcome of this in MIA is <j> orthographically (i.e. /ʤ/ in present-day Hindi), which, as Masica (1991: 169) notes, had "at least a fricative pronunciation in the Early MIA period": e.g. Skt. /jɑʋɑ/ > Pāli /ʒɑʋɑ/ 'barley'; Skt. /juddʰɑ/ > Pāli /ʒuddʰɑ/ 'battle, war'.[14] Furthermore, word-medial stop gemination before /j/ had a similar effect on sonority: e.g. Skt. /drɑʋjɑ/ > Pāli /dɑbbɑ/ 'property'. In examples like these, the high-sonority /j/-onset in the Sanskrit form was eliminated by gemination and hardening of syllable-final /ʋ/ and elimination of the /j/ in a way that is reminiscent of the well-known case of West Germanic gemination (e.g. Proto-Germanic [bid.jan] > Old English [bid.dan] 'to ask').

The final development at Stage 6 in Figure 2 involves a reanalysis of /l/-epenthesis. As was also the case with lexicalisation of /-ʋɑ/ in the transition from MIA to NIA, Stage 6 represents a change in status for causative /-lɑ/. Whereas as /l/-causatives are generated by the phonology at Stage 5 through epenthesis of /l/ to VV-sequences that occur because of suffixation of causative /-ɑ/ to vowel-final bases, Figure 2 assumes that this pattern entered a new phase in its life cycle at Stage 6. The outcome of this is a lexical listing of the first causative suffix allomorph in which /l/ is present underlyingly.[15]

[14] In some cases, preservation or re-introduction of Sanskrit forms has created doubles: e.g. Hindi /ʤʊʤʰ/ from MIA /ʒuddʰɑ/, which exists alongside /jʊdʰ/ from Skt. /juddʰɑ/.

[15] Relatedly, see Baronian & Royer-Artuso (2024 [this volume]) regarding the claim that schwa epenthesis in Armenian may, in some cases, be understood as the result of lexical restructuring, i.e. lexicalisation of forms containing underlying schwas.

As already alluded to, lexicalisation of causative /-lɑ/ as an allomorph of /-ɑ/ in the final stage of the life cycle represents what is probably the best characterisation of these patterns synchronically. The schematisation shown in Figure 4 illustrates causative-suffix *selection* based on the phonological shape of the base.

	a. /sʊn/ 'hear'	b. /so/ 'sleep'
input:	sʊn + { ɑ / lɑ }	so + { ɑ / lɑ }
	⇓	⇓
output:	[sʊn**ɑ**]	[sʊ**lɑ**]

Figure 4: Synchronic derivation of first causative stems

In Figure 4a, the /-ɑ/ causative allomorph is selected because the root /sʊn/ is consonant-final. This yields the well-formed causative stem, [sʊnɑ], in contrast to ill-formed *[sʊnlɑ]. However, affixation of the /-ɑ/ allomorph in Figure 4b would generate an ill-formed output with hiatus, i.e. *[sʊɑ]. Since the root /so/ is vowel-final, selection of the /-lɑ/ allomorph yields an optimal output with /l/-insertion, i.e. [sʊlɑ].

As shown in Figure 5, these operations can be derived from the constraint ranking already established in Figure 3b. For input Figure 4b, in which the /-ɑ/ suffix allomorph is selected – i.e. Figure 5i–iii – all candidates are eliminated by the top-ranked markedness constraints. Conversely, candidate Figure 5vi, in which the /-lɑ/ allomorph is selected and faithfully mapped to the surface form,[16] [sʊlɑ], incurs no violations of the superordinate constraints.

/so+{-ɑ, -lɑ}/	*HIATUS	*ONS/j	DEP-seg$_{[+cons]}$	*ONS/l
i. /so+ɑ/ → [sʊɑ]	*!			
ii. /so+ɑ/ → [sʊjɑ]		*!		
iii. /so+ɑ/ → [sʊlɑ]			*!	*
iv. /so+lɑ/ → [sʊɑ]	*!			
v. /so+lɑ/ → [sʊjɑ]		*!		
☞ vi. /so+lɑ/ → [sʊlɑ]				*

Figure 5: Stage 6: Suffix allomorph selection in /so+{-ɑ, -lɑ}/

[16] In this connection, see Uffmann (2024 [this volume]) for discussion of epenthesis as an outcome of faithfulness operations.

4.3 Some residual issues

The foregoing analysis has argued in favour of a pattern of causative-stem formation in synchronic Hindi in which a lexically listed suffix allomorph, /-lɑ/, is selected for vowel-final bases as this maximally satisfies constraints penalising hiatus and the generation of forms with syllable-initial /j/. This sketch analysis illustrates how such forms are derived synchronically under the assumption that a historical *phonological* pattern of /l/-epenthesis in causative formations was reanalysed diachronically as allomorphy. In this connection, there are other factors that merit comment, not least the fact that Hindi does also exhibit pre-suffixal /j/ in masculine perfective forms, as shown in Table 9. Space does not permit full discussion of how this pattern coexists and interacts synchronically with the causative-formation patterns. However, the derivations in Table 11 below briefly lay out a possible solution to this.

Table 11: Causative and perfective formation in synchronic Hindi

	Perfective formation		Causative-perfective formation	
Root:	a. /sʊn/	b. /so/	c. /sʊn/	d. /so/
	↓	↓	↓	↓
SL:	sʊn	so	sʊn-ɑ$_{\text{CAUS}}$	sʊ-lɑ$_{\text{CAUS}}$
	↓	↓	↓	↓
WL:	sʊn-ɑ$_{\text{PERF}}$	so-jɑ$_{\text{PERF}}$	sʊnɑ-jɑ$_{\text{PERF}}$	sʊlɑ-jɑ$_{\text{PERF}}$

Table 11 assumes the same stratified phonological architecture as the life cycle (cf. Bermúdez-Otero 2011, 2017, Kiparsky 2000): the distinction between stem-level (SL) and word-level (WL) processes is particularly important here. More specifically, I assume that perfective formation as a general inflectional pattern is confined to the word level. By contrast, the formation of causative stems is handled by the stem-level grammar. Thus, in examples (a) and (b), the underlying forms of the verbal stems /sʊn/ and /so/ received a faithful mapping at the stem level (i.e. in the absence of other morphological material). At the word level, the perfective suffixes attach to the outputs generated at the stem level. This yields forms in [-ɑ] in cases like [sʊnɑ] 'heard' (formed on a consonant-final base) and forms in [-jɑ] with vowel-final bases, as in [sojɑ] 'slept'.

In the case of the causatives, Table 11c is similar to Table 11b. The difference is that affixation of the causative suffix – specifically the /-ɑ/ allomorph – occurs at the stem level. This generates a causative-stem output formed on the root /sʊn/ that is vowel-final, i.e. [sʊnɑ]. Perfective formation at the word level then applies in the same way as in Table 11b, thereby yielding an output in [-jɑ]: [sʊnɑjɑ] 'caused to hear'. Table 11d presents a case of double repair. Here, formation of a causative stem with /-ɑ/ at the stem level would produce an ill-formed output with hiatus, i.e. *[sʊ-ɑ]. Accordingly, the /-lɑ/ allomorph is selected instead: this generates the /l/-causative form, [sʊlɑ]. In the same way as Table 11c, this causative stem is vowel-final. Thus, the word-level grammar generates the perfective form [sʊlɑjɑ] 'put to sleep', *[sʊlɑɑ].[17]

An obvious question that arises regarding these patterns is whether the /j/ in perfective forms is epenthetic. This is not something that can be dealt with decisively here; however, the historical literature at least suggests that examples like Table 11b–d are cases of epenthesis. Masica (1991: 269) comments that the perfective suffixes developed from a productive late Sanskrit suffix, /-itɑ/. Then, "[b]y the regular process of phonological attrition in MIA and NIA this became *-ia* and thence *-i* or *-y*. [...] The extreme weakness of this element made it prone to disappear entirely [...]. In Standard Hindi [...] it is retained only with vowel stems: *khā-y-ā* 'eaten', *gā-y-ā* 'gone'". Similarly, Mayrhofer (1951: §439) refers to a *Bindevokal* that developed from diachronic weakening of /-itɑ/. However, if it is the case that /j/ in perfective forms is epenthetic, it remains unclear why it surfaces only in the context of /-ɑ/ (i.e. in masculine forms) and not preceding other inflectional vowels (cf. the data in Table 9).

Despite this, the development of the perfective suffixes from /-itɑ/ is important in connection with the proposed historical reconstruction shown in Figure 2. As this assumes that /l/-causatives developed from historical /j/-epenthesis, a relative chronology of the related patterns is suggested. This is outlined in (2).

(2) Relative chronology of /j/-related patterns
  a. Elimination of /j/-onsets through fortition and gemination in MIA.
  b. Emergence of /j/-epenthesis in causative formations.
  c. Change from /j/ > /l/-epenthesis as sonority optimisation.
  d. Lexicalisation of /-lɑ/ as an allomorph of causative /-ɑ/.
  e. Emergence of /j/-epenthesis in perfective formations.

As discussed, the elimination of /j/ through hardening and gemination were early patterns already visible in MIA. Further to this, if the the re-introduction of /j/ in causative forms through historical epenthesis did indeed occur, then this, and the subsequent lateralisation of the epenthetic /j/, must also have been

[17] Assuming the derivation proceeds in this way also accounts for the opaque patterning of the surface vowels in perfective forms like [sojɑ] vs [sʊlɑjɑ].

relatively early developments. Under this scenario, a second, later innovation then resulted in /j/-epenthesis in the context of *perfective* /-ɑ/ via progressive weakening of the /-itɑ/ suffix over time. Crucially, the suggestion here is that this later pattern of /j/-epenthesis in perfectives parallels the earlier development which targeted *causative* /-ɑ/.[18]

To close the discussion, it must be reiterated that some of these suggestions are tentative, and the historical developments summarised in Figure 2 and the chronology in (2) constitute an attempt at synthesising various pieces of historical evidence, many of which are not fully understood independently. In addition to these historical concerns, there is also certainly more to say about the synchronic interaction between /l/-causatives and /j/-perfectives. However, this is something that must be left for now for future research to address.

5 Conclusion

The goal of this paper was to discuss the diachrony of lateral epenthesis, which, as a rare phonological phenomenon, has received little attention in the theoretical literature to date. I have argued that the development of /l/-epenthesis in English dialects is best understood as the result of rule inversion, which is consistent with previous accounts and those of the closely related phenomenon of /r/-sandhi. In Oceanic, initial /j/-epenthesis in Motu and Fijian had phonological knock-on effects. The development of initial /l/ in Motu and /c/ in Fijian has been explained in relation to language-specific optimisation of onset sonority.

Regarding Indo-Aryan, I have argued that the origin of /l/-insertion in Hindi first causative formations cannot reliably be traced – at least directly – to OIA /p/-causatives. Other explanations based on analogy from *pālayati* or /l/ in participial forms that occur in Eastern NIA languages are similarly unsatisfactory. An alternative pathway of change that does not assume an analogical development has therefore been proposed. This reconstruction assumes a *language-internal* development that involves the emergence of a pattern of /j/-epenthesis in causative forms and a subsequent change from epenthetic /j/ to /l/, paralleling the evolution of epenthesis in Motu. The proposed reanalysis has therefore aimed to account for the facts in a way that is informed by the core principles of the life cycle and that does not rely exclusively on historical dialectal patterns that may well have had no influence on the development of /l/-causatives in Hindi.

[18] Recall Blust's claim that /j/-epenthesis in Fijian was a gradual change that operated in identifiable phases.

Acknowledgements

A previous version of the analysis of Hindi causative verbs was presented at the *Fifth Edinburgh Symposium on Historical Phonology* in December 2021, in addition to the *Epenthesis and Beyond Workshop*. I am grateful to attendees at both meetings who provided feedback on the developing analysis. I would like to personally thank Nimisha Upadhyay and Aarushi Asthana who kindly provided judgements on some of the Hindi data. Thanks are also due to the two anonymous reviewers who commented on a previous draft of this chapter and offered helpful suggestions. I am responsible for any remaining errors.

Whilst preparing this paper, I was saddened to hear of the deaths in early 2022 of both Robert Blust and Colin Masica. Although I never had the opportunity to meet them, I wish to acknowledge the extraordinary contributions that both made to their respective fields of expertise. It will be immediately apparent that this paper would not have been possible without their ground-breaking work.

References

Alba, Orlando. 1979. Análisis fonológico de /r/ y /l/ implosivas en un dialecto rural dominicano. *Boletín de la Academia Puertorriqueña de la lengua española* 7. 1–18.

Anderson, Stephen R. 1981. Why phonology isn't natural. *Linguistic Inquiry* 12(4). 493–539.

Baronian, Luc & Nicolas Royer-Artuso. 2024. Textsetting the case for epenthesis in Armenian. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 271–289. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264550.

Barras, William. 2011. *Sociophonology of rhoticity and r-sandhi*. PhD thesis, University of Edinburgh.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2019–2048. Malden: Wiley-Blackwell.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 374–399. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2017. Stratal phonology. In S. J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge handbook of phonological theory*, 100–134. Abingdon: Routledge.

Bermúdez-Otero, Ricardo & Kersti Börjars. 2016. Markedness in phonology and in syntax: The problem of grounding. *Lingua* 116. 710–756.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 691–720. Oxford: Oxford University Press.

Bhatt, Rajesh & David Embick. 2017. Causative derivations in Hindi-Urdu. *Indian Linguistics* 78(1–2). 93–151.

Bloch, Jules. 1965. *Indo-Aryan: From the Vedas to modern times [English edition]*. Paris: Librairie d'Amérique et d'Orient Adrien-Maisonneuve, translated by Alfred Master.

Blust, Robert. 1999. Three recurrent changes in Oceanic languages. In Jeremy H. C. S. Davidson (ed.), *Pacific island languages: Essays in honour of g. B. Milner*, 7–28. London: University of London, SOAS.

Bokhari, Hassan. 2024. The patterning of epenthesis in Urban Hijazi Arabic. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 85–104. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264534.

Bubenik, Vik. 2003. Prākrits and Apabhraṁśa. In George Cardona & Dhanesh Jain (eds.), *The Indo-Aryan languages*, 204–249. Abingdon: Routledge.

Endzelin, J. 1922. *Lettische Grammatik*. Riga: Kommissionsverlag A. Gulbis.

Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16. 29–54.

Gick, Bryan. 2002. The American intrusive L. *American Speech* 77. 167–183.

Gopal, Deepthi. 2018. *Similarity and representations in sonorant phonology*. PhD thesis, University of Manchester.

Gouskova, Maria. 2004. Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology* 21. 201–250.

Hall, T. Alan. 2013. How common is r-epenthesis? *Folia Linguistica* 47. 55–87.

Iosad, Pavel. 2020. The life cycle of preaspiration in the Gaelic languages. In J. Kopaczyk & R. McColl Millar (eds.), *Language on the move across domains and communities: Selected papers from the 12th triennial forum for Research on the Languages of Scotland and Ulster*, 200–230. Aberdeen: University of Aberdeen.

Jamison, Stephanie W. 1983. *Function and form in the -áya-formations of the Rig Veda and Atharva Veda*. Göttingen: Vandenhoeck & Ruprecht.

Johnson, Wyn & David Britain. 2007. L-vocalisation as a natural phenomenon: Explorations in sociophonology. *Language Sciences* 29. 294–315.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The linguistic review* 17. 351–367.

Krug, Manfred. 2017. The Great Vowel Shift. In Alexander Bergs & Laurel Brinton (eds.), *The history of English: Early modern English*, 241–266. Berlin: Mouton de Gruyter.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Mayrhofer, Manfred. 1951. *Handbuch des Pāli (mit Texten und Glossar): Eine Einführung in das sprachwissenschaftliche Studium des Mittelindischen*. Heidelberg: Carl Winter Universitätsverlag.

Mayrhofer, Manfred. 1978. *Sanskrit-Grammatik (mit sprachvergleichenden Erläterungen)*. Berlin: Walter de Gruyter.

Mompeán-González, José & Pilar Mompeán-Gillamón. 2009. /r/-liaison in English: An empirical study. *Cognitive Linguistics* 20. 773–776.

Morley, Rebecca L. 2015. Deletion or epenthesis? On the falsifiability of phonological universals. *Lingua* 154. 1–26.

Norman, Kenneth R. 2012. Dialect variation in Old and Middle Indo-Aryan. In George Erdosy (ed.), *The Indo-Aryans of ancient South Asia: Language, material culture and ethnicity*, 278–292. Berlin: De Gruyter.

Oberlies, Thomas. 2001. *Pāli: A grammar of the language of the Theravāda Tipiṭaka (with a concordance to Pischel's Grammatik der Prakrit-Sprachen)*. Berlin: de Gruyter.

Oberlies, Thomas. 2003. Aśokan Prakrit and Pāli. In George Cardona & Dhanesh Jain (eds.), *The Indo-Aryan languages*, 161–203. Abingdon: Routledge.

Ohala, Manjari. 1983. *Aspects of Hindi phonology*. Delhi: Motilal Banarsidass.

Ramsammy, Michael. 2015. The life cycle of phonological processes: Accounting for dialectal microtypologies. *Language and Linguistics Compass* 9(1). 33–54.

Roberts, Philip J. 2012. Latin rhotacism: A case study in the life cycle of phonological processes. *Transactions of the Philological Society* 110. 80–93.

Sande, Hannah. 2024. Insertion or deletion? CVCV/CCV alternations in Kru languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 21–55. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264530.

Sen, Ranjan. 2016. Examining the life cycle of phonological processes: Considerations for historical research. *Papers in Historical Phonology* 1. 5–36.

Sen, Sukumar. 1960. *A comparative grammar of Middle Indo-Aryan*. Deccan College: Linguistic Society of India.

Shevelov, George Y. 1964. *A prehistory of Slavic: The historical phonology of common Slavic*. Heidelberg: Carl Winter Universitätsverlag.

Sóskuthy, Márton. 2013. Analogy in the emergence of intrusive-r in English. *English Language and Linguistics* 17. 55–84.

Trudgill, Peter. 1999. Dialect contact, dialectology and sociolinguistics. *Cuadernos de Filología Inglesa* 8. 1–8.

Turton, Danielle. 2016. Synchronic stratum-specific rates of application reflect diachronic change: Morphosyntactic conditioning of variation in English /l/-darkening. *Papers in Historical Phonology* 1. 130–165.

Turton, Danielle. 2017. Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology* 8(1). 1–31.

Uffman, Christian. 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29. 451–476.

Uffmann, Christian. 2024. Epenthesis as a matter of FAITH. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 123–141. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264538.

Vaux, Bert. 2001. *Consonant epenthesis and hypercorrection*. Paper presented at the LSA, Washington, D.C.

Vennemann, Theo. 1972. Rule inversion. *Lingua* 29. 209–242.

Weissmann, E. 1970. Phonematische Analyse des Stadtdialektes von Bristol: II. Teil. *Phonetica* 21. 211–240.

Chapter 12

Textsetting the case for epenthesis in Armenian

Luc Baronian[a] & Nicolas Royer-Artuso[b]

[a]UQAC [b]UQAC/EFMI

The authors analyze the textsetting of an Armenian song (*Ooska gukas*, recorded by the Gomidas Band in Philadelphia on Roulette Records, 1963) that shows schwas within words where they are unexpected when compared to the standard language and known dialects of Armenian. The beat of the song is a 10/16 djurdjuna in Ottoman music. The authors demonstrate that schwa epenthesis is used by the singer (Roger Mgrdichian 1930–2019) as a the main strategy to fill in beats with additional syllables in the textsetting process, within certain consonantal contexts. Besides the intrinsic interest of the textsetting process in an Armenian dialect that is now nearly extinct, this case study strengthens the point of view that schwa epenthesis is an active and productive process in Armenian, suggesting also that moraicity plays a role in the language's prosody. This is not to say that some Armenian schwas cannot be lexicalized or morphologized, but that epenthesis is live enough to be used in creative ways by speakers when playing with language.

In memory of Roger Mgrdichian, 1930–2019.

1 Introduction

In this paper, we present a case study from textsetting that offers one argument for the synchronic status of schwa epenthesis in Armenian. We do not claim to have definitely proven that schwa epenthesis is active in all varieties of Armenian for every speaker; we offer a case study of a 20th century Armenian diaspora speaker who used schwa epenthesis productively in the process of textsetting his lyrics to the (djurdjuna) beat of a song. Analyzing such a process in a language

Luc Baronian & Nicolas Royer-Artuso. 2024. Textsetting the case for epenthesis in Armenian. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 271–289. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264550

other than English requires a lengthy enough exposition that we feel makes our contribution worthwhile, even though it is only one brick in the enterprise to validate the synchronic status of schwa epenthesis in Armenian. In the general view of epenthesis as a prosodic phenomenon, we feel also that this contribution provides a new tool, which, to the best of our knowledge, has not been used so far for this purpose. As the reader will notice in most contributions to this volume, but especially in the papers by Hall (2024 [this volume]), Krämer (2024 [this volume]), Mansfield et al. (2024 [this volume]), Nelson (2024 [this volume]), Rubin & Kaplan (2024 [this volume]) and Sande (2024 [this volume]), a discussion of epenthesis without mentioning linguistic prosody is almost impossible. In this sense, we thought that the interaction of epenthesis with musical and rhythmical prosody offered something different but in line with the other contributions.

What lead us to delve into this problem is the assumption by Vaux (1998) that schwa epenthesis is a completely systematic and synchronic process, while Baronian (2017) adopts a more nuanced view. He analyzes cases where the presence of schwas must be part of the underlying form in Modern Western Armenian, even, in some cases, where there historically was an epenthesis. Even Baronian, however, still treats most schwas as synchronically epenthetic. The question behind this paper is thus: given that at least some schwas can be analyzed as part of the underlying representation, is the synchronic status of this process still valid in Modern Armenian? We intend to convince the reader that the textsetting of an Armenian folk song that we analyzed argues for considering Armenian schwa epenthesis as a still productive synchronic phonological process in the language, living next to underlying schwa, which itself can even sometimes be the result of a historical epenthesis. In doing so, we highlight the ties of epenthesis as a phenomenon to the prosody of a language.

Epenthetic processes, while being easy to define a priori, “vary enormously in their characteristics, and many aspects of their typology are still not well understood” (Hall 2011: 1). As in the case of any phonological process that linguists analyze, frameworks and/or theories will often guide the solution adopted. Competing frameworks/theories will generally offer one of two different solutions and it is sometimes difficult to decide between them. In the present case, the possibilities are: 1) phonological insertion of a schwa by an epenthetic process during the derivation from underlying representations (UR) to surface representation (SR); or 2) presence of the schwa in the underlying representation. We find it useful to search for methods and evidence external to – one might say neutral from – strictly linguistic frameworks in order to better understand a specific process.

The "external" evidence that we propose to use is the process of textsetting and its formal analysis in Generative Metrics/Generative Textsetting. That is, roughly, the analysis of the way poets and/or songwriters put their words onto metrical grids when they compose poems and songs. For example, we can see in (1) that in the English iambic pentameter, Shakespeare aligns stressed syllables (in bold) with the strong positions (s) of superimposed iambs and stressless syllables with weak positions (w):[1]

(1)	w	s	|	w	s	|	w	s	|	w	s	|	w	s
	But	thy		e-	**ter-**		nal	**sum-**		mer	**shall**		not	**fade**,

As De Sisto (2020: 1) puts it:

> The characteristics of poetic metre recreate what is attested in the phonology of the language in which verse is written (Kiparsky 1973, Hayes 1989, Fabb 1997, Golston & Riad 1999). Metre is, therefore, an abstract structure which is constructed by mirroring phonological structure and which is filled by phonological material.

The rationale of our argument is therefore the following:

1. Some processes can receive different analyses depending on the model/theory we work with;
2. Some constraints imposed by textsetting that are not part of the constraints of the language might activate such a process (e.g. by creating a different kind of context, or by creating rare or unexpected structures);
3. The speaker's reaction to this new type of context can allow us to better understand what this process is, thereby giving us some cues about the competence of the native speakers with regards to this process.

2 Aims of the contribution

Our contribution has two main goals:

1. The first goal concerns method: we want to show how using textsetting as data can help the phonologist decide if a phonological element is underlyingly represented (or not), and thus, if a phonological process – in the present case, *epenthesis* – is or is not involved in the surface variation that we observe in our data.

[1]Function words can be placed in either strong or weak positions.

2. The second goal is specific to Western Armenian: we want to help advance the answer to an important question in the phonology of Armenian, namely whether some schwas in this language are truly epenthesized synchronically.

3 About Armenian and Western vs. Eastern dialects

Armenian has its own alphabet, which, tradition holds, was invented in the 5th century for Classical Armenian by a monk named Mesrop. Classical Armenian or Grabar 'the written word' is the oldest attested written variety of Armenian. The alphabet continues to be used for the two modern standards, Western and Eastern Armenian, but has also been used for many dialects of the language and even for several other languages spoken by Armenians, most famously Ottoman Turkish up until the beginning of the 20th century.[2] In the 20th century, there was a series of spelling reforms in Soviet Armenia (Dum-Tragut 2009), which affected the spelling of words for Eastern Armenian in Armenia, but not for Eastern Armenian communities in Iran, nor for Western Armenian in the diaspora. The song we analyze in this paper is sung in a non standard dialect that belongs to the Western group, which consists mostly of dialects once spoken in the Ottoman Empire. As a result, Standard Western Armenian and Western Armenian dialects are mainly spoken by descendants of the 1915 Armenian genocide survivors.

Armenian reformed spelling is rather phonetic. The more traditional spelling still favored by Western Armenian speakers has, as one would expect, a greater discrepancy with pronunciation, but is still much closer to pronunciation than English or French might be to their respective spelling systems. One example of a discrepancy is that final Yi (Յ/յ) is often silent word-finally in the traditional orthography, representing a former 3sg suffix once pronounced [j] or sometimes part of a case suffix[3] once pronounced [-ɑj], but now pronounced [-ɑ]:

(2)	Traditional spelling	Կարդայ	'He reads'
	Reformed spelling	Կարդա	
	Classical Armenian SR	[karday]	
	Standard Eastern Armenian SR	[kɑɾdɑ]	
	Standard Western Armenian SR	[gɑɾtʰɑ]	

[2]For example, the US Library of Congress holds in its catalog several 19th century texts in what it terms Armeno-Turkish.

[3]The suffixed form is genitive, possessive or ablative. In the song analyzed, it is used once as an ablative.

These silent letters were removed in the reformed spelling. Another example of discrepancy is that Yi (Յ/յ), Vo (Ո/ո) and Ech (Ե/ե) are respectively pronounced [h], [vɔ] and [jɛ] word-initially in the traditional orthography, but [j], [ɔ] and [ɛ] word-medially. Initial Yi was thus replaced by Ho (Հ) in the reformed spelling, but the other two letters have been preserved in this position even in the reformed spelling.

(3)

Traditional spelling	Յակոբ	'Jacob'
Reformed spelling	Հակոբ	
Classical Armenian SR	[jɑkob]	
Standard Eastern Armenian SR	[hɑkob]	
Standard Western Armenian SR	[hɑgop^h]	

One discrepancy that is specific to Standard Western Armenian and some dialects is the merger of two series of stop consonants and affricates: Standard Western Armenian opposes voiceless aspirates to voiced stops and affricates, while Standard Eastern Armenian and more conservative dialects have a trilateral opposition, usually[4] between voiced, voiceless unaspirated and voiceless aspirated stops and affricates. In this case, making orthography correspond to pronunciation would involve removing five letters from the alphabet, which is not likely to be viewed favorably by most Armenians. The dialect used in the song under study in this paper, however, has a different merger than Standard Western Armenian as illustrated in (4).

(4) Illustration of the stop and affricate mergers in Standard Western Armenian and in the dialect used in the song under study

	'petal'	'still'	'father (addressing a priest)'
Traditional spelling	Թեր	Դեռ	Տէր
Reformed spelling	Թեր	Դեռ	Տեր
Standard Eastern SR	[tʰɛɾ]	[dɛr]	[tɛɾ]
Standard Western SR	[tʰɛɾ]	[tʰɛɾ]	[dɛɾ]
Dialect of the song SR	[tʰɛɾ]	[dɛɾ]	[dɛɾ]

Another relevant feature of the alphabet for this paper is the fact that there exists an Armenian letter for schwa, called Et (Ը/ը). The rule of thumb in both the traditional and reformed spellings is that this letter is used whenever it is not predictable by epenthesis. This has given phonologists a handle on deciding which

[4]Pisowicz (1976) lists a total of seven voicing patterns. Some dialects have voiced aspirated stops and affricates, sometimes termed murmured. For details, see Baronian (2017). For the phonetic nature of voiced aspirates, see Khachaturian (1992), Seyfartha & Garellek (2018).

schwas are part of the UR and which schwas are derived through epenthesis. As we will see, some phonologists tend to posit less UR schwas than orthography suggests, but there is also nothing preventing us from positing more schwas than orthography suggests if we assume that orthography is generally more conservative than the spoken language.

4 The phonological problem: Underlying schwa vs schwa-epenthesis in Western Armenian

As highlighted by Baronian (2017), some Armenian schwas can be analyzed as part of the underlying representation. For example, the definiteness or specificity suffix of Western Armenian (Sigler 1996) makes consonant-final nouns alternate with schwa-less forms, whereas vowel-final nouns use the allomorph /-n/:

(5)	T/RS[5]	մատ	մատը	լեզու	լեզուն
	WT	mad	madë	lezun	lezu
	U/SR	mɑd	mɑdə	lɛzu	lɛzun
		‘finger’	‘the finger’	‘tongue’	‘the tongue’

Vaux (1998)’s earlier analysis posited a unified suffix *-n* that triggered epenthesis when forming a cluster and a special rule that deleted the *-n* later (therefore in the example above, underlying /mad-n/ would become [madən] before resulting in surface [madə]). We favor the less abstract allomorphic analysis, because the n-deletion rule proposed by Vaux, while it almost certainly corresponds to what happened historically, does not appear to have survived elsewhere in the language.[6] In our view, positing a suffix-special rule does not place any less burden on memory than positing V/C-sensitive allomorphs. Therefore, minimizing the level of abstraction in the derivation should be favored by Occam’s razor.

[5]Following comments by two anonymous reviewers, we tried to use either TS (traditional spelling) or RS (reformed spelling) based on which we thought was most useful to help the reader understand how we determined the UR. Our transliteration system WT is basically the ISO 9985 romanization system for Armenian, with two exceptions: 1) we transliterate the digraph ու as u instead of ow, because this digraph always represented a single vowel; 2) we switched to Western Armenian values for unaspirated stops and affricates, because the dialect is Western and not doing so would have distracted the reader from voicing issues not relevant to the question of epenthesis.

[6]Except, as Baronian (2017) points out, before the verb for ‘be’ and before the word *al* ‘also’.

(6)		T/RS	մատը	
		WT	madë	
	Vaux's analysis			Baronian's analysis
	UR mɑd-n			UR mɑd-ə
		SR	mɑd-ə	

Whatever one's view on the definiteness or specificity suffix is, the schwas that have attracted most attention in phonology are those that interrupt a consonant sequence otherwise unattested in Armenian and assumed to be impossible to pronounce by native speakers. Examples from Western Armenian are given in (7):

(7)	T/RS	նկար	պտտիլ
	WT	ngar	bddil
	UR	ngɑr	bddil
	SR	nəgɑr	bədədil
		'portrait'	'to stroll'
		*[ng...] unattested	*[bdd...] unattested

As Baronian (2017) points out, if we assume a sonority hierarchy Stops < Fricatives < Nasals < Liquids < Glides < Vowels, only the onset C-glide clusters and some of the C-liquid clusters are allowed to remain as such in the SR. The other onsets (whether sonority rises or falls) break up the cluster by inserting a schwa, except in the context of sC, where epenthesis precedes the cluster, as it does in Modern Spanish or Old French, for example.[7] In codas, only clusters with raising sonority are broken up by epenthesis:

(8)	T/RS	վագր	Ագն
	WT	vakr	agn
	UR	vɑkr	ɑgn
	SR	vɑkər	ɑgən
		'tiger'	town's name
		*[...kr] coda unattested	*[...gn] coda unattested

Epenthesis can also apply in some codas with falling sonority, but only when they contain the possessive suffix -*s* (1sg) or -*t* (2pl). In this case, as recognized by Baronian (2017), an analysis that would posit lexicalized -*əs* and -*ət* as allomorphs

[7]For example, Armenian Ստեփան /sdepan/ 'Stephen' is pronounced [əsdep^{h}ɑn], similar to the Spanish cognate Esteban and the French cognate Etienne (Old French Estienne).

is possible, though he favors still considering the suffixes to trigger epenthesis. We will return to this special but crucial case after analyzing the song.

Etymologically speaking, it can be shown that the schwas in (7) and (8) were epenthesized at some point, but one may still wonder how to prove the synchronic status of their epenthesis. The fact that they are spelled without schwas in Armenian orthography should not automatically make us conclude that they are not part of the UR, even though Armenian orthography is closer to pronunciation than English or French is to their respective written forms.

Because one never hears the root [bədəd-] without its schwas, it is certainly possible that at least some speakers lexicalize it as /bədəd-/ instead of /bdd-/ suggested by the orthography, even though the schwas are entirely predictable from the way epenthesis works in this language. In fact, in the case of onset C-liquid clusters, it is probably the case that they were forbidden historically, resulting in /grɑg/ 'fire' being pronounced [gərɑg], but that this requirement was laxed for Modern Armenian in some traditional words, resulting in /krikor/ 'Gregory' being pronounced [krikor], and in new borrowings like *Gloria* or *iCloud*. It is then probably simpler to consider the UR for 'fire' to be /gərɑg/ in Modern Armenian, and let epenthesis apply only in onsets where the second consonant is less sonorous than liquids. A more radical approach might propose that onset epenthesis has disappeared altogether from the language.

For example, in a case like /ngɑr/ pronounced [nəgɑr], there is even a near-minimal pair with the word pronounced [ənger] meaning 'friend', where the schwa is placed differently. In this case, orthography encodes the schwa in the latter (T/RS ընկեր, WT ënger), but not in the former (T/RS նկար, WT ngar), suggesting epenthesis is active in [nəgɑr], but not in [ənger]. However, these orthographic choices do not prove anything: what is to prevent a speaker from lexicalizing both forms, each with a schwa in a different location?

What these examples show us is that there clearly was an active epenthesis in the language historically, but its conditions have been laxed over time, to a point where we may wonder whether epenthesis is still synchronically active at all.

5 The song under study and its dialect

The song we found where schwa epenthesis is used productively in the textsetting process is titled *Ooska gukas* and was recorded by the Gomidas Band in Philadelphia in 1963 on Roulette Records. The singer, Roger Mgrdichian, was born in 1930 in the US, the son of Ottoman Armenian immigrants from Peri, which is now known as Akpazar in Turkey.[8]

We noticed that the dialect to which the song belongs has two defining isoglosses: 1) the merger of Proto-Indo-European (PIE) stop series II and III into a single voiced series (Type III voicing, as classified for example by Pisowicz 1976); 2) the use of *gu-* as a present marker. Both the voiced nature of PIE stop series II and the use of *gu-* as a present marker marks the dialect as unambiguously part of the Western Armenian group of dialects.

The lyrics of the song include a few Turkish borrowings, but remain largely intelligible to a Standard Western Armenian speaker. Because one of the main differences between Standard Eastern and Standard Western Armenian involves the voicing of stops and affricates, the voicing pattern of the stops in this dialect is probably the most challenging feature in terms of mutual intelligibility, as the merger, in effect, makes some words sound Eastern and some words sound Western. To the best of our assessment, however, nothing in this dialect bears on the schwas in the language, nor on epenthesis. For this reason and because the singer was part of a larger Western-Armenian speaking diasporan community in the Philadelphia area, it seemed to us that whatever conclusion we might draw about the dialect used in the song and its singer can be extended to Western Armenian speakers generally without any further assumptions.

The complexity of the situation, with the existence of two standards, multiple dialects and a spelling reform not universally used by all Armenians, made us hesitate on the question of how to best represent the words of the song. Devising dialect-specific spellings to accommodate for the differences between the Western standard and the dialect under study would not easily be recognizable by someone who reads Armenian and would not have brought any additional information to readers unable to read Armenian. Phonetic transcriptions with glosses seemed like the best avenue in such circumstances. At the same time, Armenian orthography, even the traditional one, is close to being phonemic, thereby approximating the UR, so, at the suggestion of a reviewer, we decided that providing the spellings in the annex, where the entire song can be consulted, with a general Western Armenian transliteration (which makes distinctions in voicing/aspiration that the dialect does not make), along with our phonetic transcriptions was the best option. This lessens the burden on the reader in the examples

[8] An anonymous reviewer questions the fluency level of the US-born singer. Knowing the demographics of Armenians in Philadelphia in those years, it sounds extremely unlikely to us that the singer would not have spoken Armenian fluently. In a correspondence with his son (also Roger), Baronian was able to confirm that Mgrdichian spoke Armenian and some Turkish at home. In fact, his son reports a family story, according to which his father was once sent home from the 1st grade, because he could not speak enough English.

provided to make the argument, but the reader who wants to see how we determined the UR can consult the annex. Readers should keep in mind these two points:

1. The dialect under study merges two series of stops and affricates, which, in the Classical language, were unaspirated voiceless and voiced, into a voiced series. (The Classical voiceless aspirates remain as such.)
2. As we pointed out earlier, while schwa has its own letter in the Armenian alphabet, it is not always present in the written form, cf. Baronian (2017) and similar distinctions in Dutch by van Oostendorp (2011). It is plausible that a good first approximation would state that schwa is written when it is part of the lexical form and not written when it is epenthesized synchronically, but, for the purpose of our analysis, it was important to transcribe every schwa that was present in the surface representation.

6 Data and analysis

The rhythm used in the song, called *djurdjuna*, is generally analyzed as a 10/16.[9] In (9), the strong beats are marked with an upper case X, the weak beats are marked with a lower case x, and the non-obligatory strong beats are marked within parentheses (X):

(9) Xx(X) Xx Xx Xx(X) | Xx(X) Xx Xx Xx(X)
SR dɛr[ə]dəd jɛm jɛː | ɣɛr[ə] di vɑ nɑ

SR dɛr[ə]d+əd jɛm jɛːɣɛr[ə] divɑnɑ
worry+2SG am become+Inf crazy
'I am your problem, I go crazy'

Except in the chorus, the singer seems to structure each line with two measures as follows:

(10) XxX Xx Xx Xxx | XxX Xx Xx Xxx

The textsetting rules we have identified are the following, and they were always followed:

[9]One reviewer questions how we determined the beat of the song. We do not have the space to get into musical details, but Royer-Artuso is a professional musician who has studied and practiced Ottoman music for decades. There is no doubt to either of us that this beat is recognizable as a *djurdjuna*.

(11) a. XxX must be filled by two syllables,
b. Xx must be filled by a single syllable,
c. Xxx is filled by a heavy syllable (with coda or long vowel) in the middle of a line, but can be filled by a light syllable at the end of a line.

In the sample in (9), as well as in the full text of the song provided in the Appendix in Section 8, the schwas indicated in [square brackets] are the interesting ones not expected in the pronunciation of Armenian because: 1) they are not part of the historical forms of those words; 2) they do not represent the definiteness/specificity suffix; and 3) they do not break up unattested clusters of Armenian. In fact, some very similar clusters sometimes appear elsewhere in the same song: /dɛrdd/ in (9) and line [2] of the Appendix becomes [dɛrədəd], but /ɑrnim/ (line [7] of the Appendix) remains schwaless. Our explanation for the insertion of those non-standard schwas is that the singer or composer syllabifies a consonantal mora in order to occupy a strong metrical position. The singer always inserts these "new" schwas after a coda, never before a word-initial onset or never to break up a cluster that is not already broken up by regular syllabification or epenthesis:

(12) Textsetting schwas inserted by the singer in his performance (lines in [square brackets]):

[1] us[ə]gɛ	[2] dɛr[ə]dəd	[2] jɛːɣɛr[ə]	[3] bɑtʃig[ə]
'from where'	'your worry'	'it seems'	'kiss'
[4] jɑr[ə]	[5] jɛs[ə]	[8] uʃig[ə]	[9] vod[ə]gəd
'soul'	'I'	'late-ish'	'your foot'
[9] vɑr[ə]di	[10] dur[ə]	[12] dun[ə]	[13] dur[ə]
'rose-DAT'	'give-IMP'	'you'	'door'

In most words in (12), schwa appears at the end of the word, after a single consonant. Because most of these words are not nouns, the schwa in them cannot be interpreted as a definiteness/specificity suffix. Even in *batʃig[ə]* 'kiss' and in *dur[ə]* 'door', the presence of the indefinite marker mə/mi immediately after lifts the ambiguity. In the words *us[ə]gɛ* and *var[ə]di*, schwa breaks up two consonants already belonging to two separate syllables and then does not even fall under traditional schwa epenthesis. Therefore, the schwa epenthesis studied here is most similar to the cases of epenthesis studied by Hamann & Miatto (2024 [this volume]), as well as Nelson (2024 [this volume]), as it does not break up a consonant cluster within an onset or within a coda. Contrary to Krämer's excrescent or intrusive vowel cases, however, the epenthesis here is clearly available to prosodic computation, albeit musical, not linguistic. Interestingly, we could say that the Armenian textsetting epenthesis is the mirror image of the *-a* omission mentioned in section 5.2 of the paper by Mansfield et al. (2024 [this volume]), in that both occur at the edge of a prosodic domain (the syllable for Armenian).

In the word *dɛr[ə]dəd*, the second schwa is also epenthetic, but it is expected in the regular pronunciation, because it precedes the 2SG possessive suffix *-d* (*-t* in Standard Western Armenian as discussed earlier). The unexpected epenthesis, on the other hand, breaks up the *-rd-* cluster in the sense that the two consonants now belong to different syllables. The exact same situation happens in the word *vod[ə]gəd*.

This situation is interesting, because it recalls Baronian (2017)'s analysis of the possessive suffix as being affixed with a coda status, thus revising Vaux (1998)'s analysis that posited a prosodic word boundary. The motivation for this analysis was that the suffix triggered epentheses that were not observed elsewhere in the language, thus:

(13)

T/RS	դուրս	դուռս	դուռդ
WT	turs	tuṙs	tuṙt
UR	turs	turs	turt
SR	turs	turəs	turət
	'outside'	'my door'	'your door'

In /dɛrd-d/⟶[dɛrdəd] and /vodg-d/ ⟶[vodgəd][10], epenthesis is necessary anyway, because *-rdd* or *-dgd* codas are not pronounceable in Armenian without it. However, adopting Baronian (2017)'s analysis of the possessive suffix as a consonant specified for moraicity (with the assumption that codas are moraic) offers a potential bridge to understanding the singer's strategy in the song. If epenthesis allows one to preserve the moraic nature of the consonant in the context of the possessive suffix, the singer may have simply generalized this epenthesis in order to facilitate the textsetting. More precisely, the singer syllabifies a consonantal mora by epenthesizing after the moraic (coda) consonant, whereas in the regular language, epenthesis precedes a moraic (coda) suffix within a cluster.[11] We assume that the coda position is moraic for all consonants and therefore the derivation is the following:

[10]In example (13), we used the Standard Western Armenian form of the 2SG possessive suffix (-t), but in *dɛrdəd* and *vodgəd*, which are words taken from the song under analysis, we used the dialectal form *-d*.

[11]Baronian (2017) lists two other cases of morphological operations that need to refer to syllable structure in Western Armenian.

(14)

UR	/dɛrd-d/			/vodg-d/		
Regular epenthesis	dɛrdəd			vodgəd		
Mapping to beat	X	x	x	X	x	x
	dɛ	r$	dəd	vo	d$	gəd
Textsetting epenthesis	dɛ$	rə$	dəd	vo$	də$	gəd

A legitimate question to ask at this point is whether other linguistic phenomena active in the language are used by the singer to fit the meter of the song. We believe this to be the case. For example, reduplication is active in Armenian (Vaux 1998) and the singer reduplicated the last syllable of *bəzdig* 'small' on line [10] (see Appendix in Section 8). On line [2] and in the last two lines of the song, the singer also lengthens a vowel in *jɛːɣɛr[ə]*, *divanaː* and *mɛziː*. While vowel lengthening is not a phonological process reported for Armenian, it certainly exists as an emphasis strategy. In both these cases (reduplication and lengthening), it is interesting to note that epenthesis would not have been quite applicable because, in *bəzdig*, epenthesis after [g] would not have been enough to fit the Xxx position (the final syllable in resulting *bəzdigə* would have been light, thereby violating textsetting principle C) and, in *jɛɣɛr*, the [ɣ] is not moraic, because it is located in an onset, thereby preventing textsetting epenthesis. It's true, however, that, in the case of *bəzdig*, the singer could have opted to epenthesize word-medially, yielding a metrically acceptable *bəzədig*.

7 Conclusion

The performance by Mgrdichian allows us to deepen our understanding of the synchronic status of schwas in Armenian. While we cannot prove on the basis of this song that the schwas illustrated in (7) are not lexicalized, the productivity of the generalized schwa epenthesis employed to fill what would otherwise remain as empty strong metrical positions suggests that schwa epenthesis was at least available to this speaker at the phonological level and reinforces the opinion defended by Vaux (1998), and more recently by Baronian (2017) and Dolatian (2021), that schwa epenthesis is a synchronically active process of Armenian. In particular, it is striking that the conditions for epenthesis used for textsetting purposes resemble those that occur in the context of possessive suffixation and reinforces Vaux (1998)'s opinion that epenthesis was active in this case, while it favors, however, Baronian (2017)'s analysis of the suffix as occupying a (moraic) coda position. While the challenge of explaining the textsetting mechanism in a language other than English took up too much space for us to broaden the scope of the paper beyond this case study, we hope that readers will also be convinced of the usefulness of textsetting as a tool to understand the nature of phonological processes. Finally, we also hope that readers will agree that the interplay of schwa epenthesis with textsetting in order to fill empty beat positions strengthens the view that schwa epenthesis can also be considered a prosodic process, rather than a strictly segmental one. Like the -*a* omission discussed by Mansfield et al. (2024 [this volume]), it occurs at the edge of a prosodic domain (the syllable). However, unlike the regular phonological epenthesis of Armenian (and unlike the cases studied in this volume by Bellik 2024, Bokhari 2024, Hall 2024, Krämer 2024, Mansfield et al. 2024, Rubin & Kaplan 2024 [this volume], and Sande 2024 [this volume]), the textsetting epenthesis does not break up a cluster within an onset or a coda, but follows a coda, a case similar to those under study by Hamann & Miatto (2024 [this volume]), as well as Nelson (2024 [this volume])).

8 Abbreviations

TS	Traditional spelling	UR	Underlying representation
RS	Reformed spelling	SR	Surface representation
WT	Western transliteration		

References

Baronian, Luc. 2017. Two problems in Armenian phonology. *Language & Linguistics Compass* 11. e12247.

Bellik, Jennifer. 2024. Gestural characteristics of vowel intrusion in Turkish onset clusters: An ultrasound study. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 143–166. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264540.

Bokhari, Hassan. 2024. The patterning of epenthesis in Urban Hijazi Arabic. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 85–104. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264534.

De Sisto, Mirella. 2020. *The interaction between phonology and meter*. Utrecht: LOT.

Dolatian, Hossep. 2021. *Orthography to phonology: Directional syllabification and epenthesis of the Armenian schwa*. Manuscript.

Dum-Tragut, Jasmine. 2009. *Modern Eastern Armenian*. Number 14 in London Oriental and African Language Library. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/loall.14.

Fabb, Nigel. 1997. *Linguistics and literature: Language in the verbal arts of the world.* Oxford: Blackwell.

Golston, Chris & Tomas Riad. 1999. The phonology of Classical Greek meter. *Linguistics* 38. 99–167.

Hall, Nancy. 2011. Vowel epenthesis. In Marc van Oostendorp, Collin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1576–1596. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444335262.wbctp0067.

Hall, Nancy. 2024. Intrusive and epenthetic vowels revisited. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 167–197. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264542.

Hamann, Silke & Veronica Miatto. 2024. Three language-specific phonological interpretations of release bursts and short vowel-like formants. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 199–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264544.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Paul Kiparsky & Gilbert Youmans (eds.), *Rhythm and meter*, 201–260. Orlando: Academic Press.

Khachaturian, A. 1992. Voiced aspirated consonants in the Nor Bayazet dialect of Armenian. In J. A. C. Greppin (ed.), *Proceedings of the 4th International Conference on Armenian Linguistics*, 115–128. Delmar: Caravan.

Kiparsky, Paul. 1973. The role of linguistics in a theory of poetry. *Dædalus* 102(3). 231–244. https://www.jstor.org/stable/20024157.

Krämer, Martin. 2024. Prokaryotic syllables and excrescent vowels in two Yuman languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 225–245. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264546.

Mansfield, John, Rosey Billington & Hywel Stoakes. 2024. Vowel predictability and omission in Anindilyakwa. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 57–84. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264532.

Nelson, Brett C. 2024. Insertion of [spread glottis] at the right edge of words in Kaqchikel. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 291–327. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264552.

Pisowicz, A. 1976. *Le développement du consonantisme arménien.* Wrocaw: Zakod.

Rubin, Edward J. & Aaron Kaplan. 2024. Segmental and prosodic influences on Bolognese epenthesis. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 105–121. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264536.

Sande, Hannah. 2024. Insertion or deletion? CVCV/CCV alternations in Kru languages. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 21–55. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264530.

Seyfartha, Scott & Marc Garellek. 2018. Plosive voicing acoustics and voice quality in Yerevan Armenian. *Journal of Phonetics* 71. 425–450.

Sigler, Michelle. 1996. *Specificity and agreement in Standard Western Armenian*. Massachusetts Institute of Technology. (Doctoral dissertation).

van Oostendorp, Marc. 2011. *Phonological projection: A theory of feature content and prosodic structure*. Berlin & New York: Mouton de Gruyter. DOI: 10.1515/9783110809244.

Vaux, Bert. 1998. *The Phonology of Armenian*. Oxford: Clarendon Press.

Vaux, Bert. 2003. Consonant epenthesis and the problem of unnatural phonology. Yale University Linguistics Colloquium, 16 September 2003.

Ooska gukas

Line [1]

TS	Ուսկէ		կու	գաս	Վերիէ		Վանայ	
WT	Usgē		gu	kas	veriē		vanay	
	XxX	Xx	Xx	Xxx	XxX	Xx	Xx	Xxx
UR	us	gɛ #	gu #	gɑs #	vɛ ri	ɛ#	vɑ	nɑ #
SR	us[ə]	gɛ #	gu #	gɑs #	vɛ ri	ɛ#	vɑ	nɑ #
	from where		PROG	come-2SG	upper-DAT-ABL		Van-ABL	

‘From where do you come, from Upper Van?’

Line [2]

TS	Տերտդ		եմ	եղեր		տիվանա		
WT	Derdt		em	eġer		divana		
	XxX	Xx	Xx	Xxx	XxX	Xx	Xx	Xxx
UR	dɛr	dd #	jɛm #	jɛ	ɣɛr #	di	vɑ	nɑ #
SR	dɛr[ə]	dəd #	jɛm #	jɛː	ɣɛr[ə] #	di	vɑ	nɑ #
	worry-POSS-2SG		am	it seems		going crazy		

‘I am your worry, I go crazy’

Line [3]

TS	Պաչիկ		մը	տուր		Մայրը		չի	մանա	
WT	Bačig		më	dur		mayrë		či	mana	
	XxX	Xx	Xx	Xxx		XxX		Xx	Xx	Xxx
UR	bɑ tʃig #		mə #	dur #		mɑ(j)rə #		tʃi #	mɑ	nɑ #
SR	bɑ tʃig [ə] #		mə #	dur #		mɑrə #		tʃi #	mɑ	nɑ #
	kiss		a	give-IMP-2SG		mother-DEF		NEG	watch-3SG	

‘Give me a kiss, the mother is not watching’

Chorus: Line [4]

TS	Յար		կիւլիւմ		ճան	է...			
WT	Yar		giwliwm		ǰan	ē			
	XxX		Xx	Xx	Xxx	XxX	Xx	Xx	Xxx
UR	jɑr #		gy	lym #	dʒɑn #	ɛ... #			
SR	jɑr[ə] #		gy	lym #	dʒɑn #	ɛ... #			
	Friend		rose		soul	is			

‘The friendly rose is my soul’

Line [5]

TS	Սիրեր		եմ	ես		զեզ...			
WT	Sirer		em	es		kez			
	XxX	Xx	Xx	Xxx		XxX	Xx	Xx	Xxx
UR	sirɛ	r #	jɛm #	jɛs #		kɛz... #			
SR	sirɛ	ːr #	jɛm #	jɛs[ə] #		kɛz... #			
	love		am	I		you			

‘I love you’

Line [6]

TS	Յար		կիւլիւմ		ճան	է...			
WT	Yar		giwliwm		ǰan	ē			
	XxX		Xx	Xx	Xxx	XxX	Xx	Xx	Xxx
UR	jɑr #		gy	lym #	dʒɑn #	ɛ... #			
SR	jɑr[ə] #		gy	lym #	dʒɑn #	ɛ... #			
	Friend		rose		soul	is			

‘The friendly rose is my soul’

Line [7]

TS	Պիտի	առնեմ		ես		զեզ…			
WT	Bidi	aṙnem		es		kez			
	XxX	Xx	Xx	Xxx		XxX	Xx	Xx	Xxx
UR	bi di #	ɑr	nim #	jɛs #		kɛz... #			
SR	bi di #	ɑr	nim #	jɛs[ə] #		kɛz... #			
	FUT	take-1SG		I		you			

‘I will take you away’

Line [8]

TS	Ուսկէ		կու	գաս		ուշիկ		մուշիկ	
WT	Usgē		gu	kas		ušig		mušig	
	XxX	Xx	Xx	Xxx		XxX	Xx	Xx	Xxx
UR	us	gɛ #	gu #	gɑs #		uʃi	g #	mu	ʃig #
SR	us[ə]	gɛ #	gu #	gɑs #		uʃi	g[ə] #	mu	ʃig #
	from where		PROG come-2SG			late-DIM		REDUP	

‘From where do you come so late?’

Line [9]

TS	Ոտքդ		մտեր		վարդի		բուշիկ	
WT	odk’t		mder		varti		pušig	
	XxX	Xx	Xx	Xxx	XxX	Xx	Xx	Xxx
UR	vod	gd #	md	ɛr #	vɑr	di #	pu	ʃig #
SR	vod[ə]	gəd #	məd	ɛr #	vɑr[ə]	di #	pu	ʃig #
	foot-POSS-2SG	enter		rose-DAT		thorn		

‘The rose’s thorn entered your foot’

Line [10]

TS	Տուր	պզտիկ		տիկ	պաչիկ		անուշիկ		
WT	Dur	bzdig		dig	bač̣ig		anušig		
	XxX	Xx	Xx	Xxx	XxX		Xx	Xx	Xxx
UR	dur #	bz	dig	dig #	bɑtʃig #		ɑ	nu	ʃig #
SR	dur[ə] #	bəz	dig	dig #	bɑtʃig #		ɑ	nu	ʃig #
	give-IMP-2SG	small		REDUP	kiss		sweet		

‘Give a little sweet kiss’

Line [11]

TS	Ուսկէ		կու	գաս		Վերիէ		Վանայ	
WT	Usgē		gu	kas		veriē		vanay	
	XxX	Xx	Xx	Xxx		XxX	Xx	Xx	Xxx
UR	us	gɛ #	gu #	gɑs #		vɛ ri	ɛ#	vɑ	nɑ #
SR	us[ə]	gɛ #	gu #	gɑs #		vɛ ri	ɛ#	vɑ	nɑ #
	from where		PROG	come-2SG		upper-DAT-ABL			Van-ABL

‘From where do you come, from Upper Van?’

Line [12]

TS	Ես	տիվանա			դուն		տիվանա		
WT	Es	divana			tun		divana		
	XxX	Xx	Xx	Xxx	XxX		Xx	Xx	Xxx
UR	jɛs[ə] #	di	vɑ	nɑː #	dun[ə] #		di	vɑ	nɑ #
SR	jɛs[ə] #	di	vɑ	nɑː #	dun[ə] #		di	vɑ	nɑ #
	I	go crazy			you		go crazy		

‘I go crazy, you go crazy’

Line [13]

TS	Աստուածը		մեզի		դուռ		մի	բանա	
WT	Asduaçë		mezi		tuṙ		mi	pana	
	XxX	Xx	Xx	Xxx	XxX		Xx	Xx	Xxx
UR	ɑsdvɑdz	ə #	mɛ	ziː #	dur[ə] #		mi #	bɑ	nɑ #
SR	ɑsdvɑdz	ə #	mɛ	ziː #	dur[ə] #		mi #	bɑ	nɑ #
	God-DEF		us-DAT		door		a	open-2SG	

‘May God open a door for us’

Chapter 13

Insertion of [spread glottis] at the right edge of words in Kaqchikel

Brett C. Nelson
University of Calgary

This paper examines a set of allophonic processes in Kaqchikel (ISO 639-3: cak) to determine the status of the laryngeal feature [spread glottis] in the language. In word final position: underlyingly plain voiceless stops surface as aspirated voiceless stops while sonorants surface as voiceless fricatives. I argue that these are the results of prosodic domain marking whereby [+spread glottis] is inserted at the right edge of every prosodic word. This is despite [spread glottis] not being active in contrasting the phonemes of Kaqchikel. Thus, I claim that a language's phonology may still manipulate features which are non-contrastive.

1 Background

In spoken languages, all speech sounds are made using air that passes through the larynx, specifically the glottis – the opening between the vocal folds in the larynx – which can be in one of several different states (see e.g. Ladefoged 1971, Ladefoged 1983). These laryngeal states determine the phonation of the speech sound. Critically, spoken languages make lexical contrasts based on these different phonation states, with many of these contrasts occuring at the segmental or phonemic level.

To account for these phonological contrasts made in the larynx, various sets of laryngeal features have been proposed (e.g. Halle & Stevens 1971, Clements 1985, Lombardi 1994, Iverson & Salmons 2007). Critical to the current study are the features of glottal width [spread glottis] and [constricted glottis], as well as the feature of [voice].

Brett C. Nelson. 2024. Insertion of [spread glottis] at the right edge of words in Kaqchikel. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 291–327. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264552

The feature [spread glottis] ([sg]), as defined by Halle & Stevens (1971), refers to the outward displacement of the vocal folds, and thus a widening of the glottis. Sounds typically represented by [+sg] include aspirated stops and voiceless fricatives Vaux (1998) (see also Ridouane 2006 and especially Avery & Idsardi 2001 for [sg]'s relation to glottal fricative /h/). This is opposed to [constricted glottis] ([cg]), which corresponds to "adduction of the arytenoid cartilages relative to the position of normal voicing" (Halle & Stevens 1971: 201–202). [+cg] typically results in ejectives, implosives, glottal stop, and creak (Fallon 2002). "Normal voicing" is called for by the feature [+voice], and refers to the continuous vibration of the vocal folds as air passes through the glottis (Lombardi 1994).

Languages may use any of these in distinguishing their segments. For example, Spanish (ISO 639-3: spa) uses [voice] in distinguishing its voiced stops /b d g/ from its voiceless stops /p t k/, while Korean (kor) stops exhibit a three-way contrast using two features: [cg] and [sg] to distinguish [+cg] [−sg] fortis stops /p' t' k'/ from [−cg] [+sg] aspirated stops /p^h t^h k^h/ from [−cg] [−sg] plain (lenis) stops /p t k/ (see Hamann & Miatto 2024 [this volume] for a discussion of Korean stop perception). Indeed it is even possible for a language to not utilize any of these laryngeal features contrastively. For example, Plains Cree (crk), has just a single series of three stops differentiated only by Place /p t k/.

Critically for the current study, laryngeal features may be implemented to strengthen a boundary between prosodic domains (Cho 2016) and, in the extreme case, inserted by the prosodic structure (Iosad 2016), in effect turning a laryngeally unmarked segment into a marked one (e.g. [+sg] turning [t] into [t^h] as is argued for final fortition in German (deu) (Iverson & Salmons 2007).

A question remains, however, in that the set of laryngeal features that are inserted for enhancement or edge-marking purposes in a given language is typically restricted to the set of laryngeal features that are contrastive in that language. Can non-contrastive features, or those not active in distinguishing phonemes, be active in the marking or enhancement of prosodic domains?

I investigate this question in the rest of this paper, beginning with the next section in which I introduce the language of study, Kaqchikel (cak), and relevant parts of its phonological systems. In §3, I present the allophony that occurs among several subsets of Kaqchikel's phonemes, specifically in different word positions. These cases of allophony lead to my proposal in §4. I then discuss this proposal in §5 in relation to alternative explanations as well as similar cases in other languages. Finally, in §6, I summarize and conclude the study.

2 Kaqchikel

2.1 Kaqchikel language background

Kaqchikel is a K'ichee'an language in the Mayan language family. It is spoken by about 400,000 first language speakers, primarily in south-central Guatemala (see Figure 1), but with speakers in diaspora across North America (Heaton & Xoyón 2016).

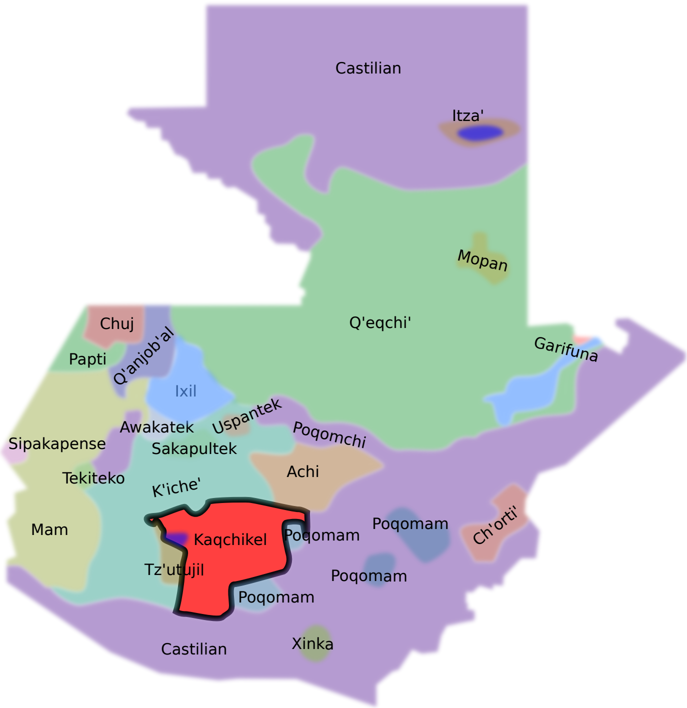


Figure 1: Map of languages of Guatemala (adapted from http://commons.wikimedia.org/wiki/File:Idiomasmap_Guatemala.svg CC BY-SA 3.0 Ignacio Icke)

Considered an at risk language by both UNESCO (Moseley 2010) and Ethnologue (Lewis 2009), it has co-official language status (with Spanish) in the regions where it is spoken (Guatemala 2003). As such, Kaqchikel people are guaranteed the right to use their language in public spheres, and to have their language developed, used, and interpreted in educational, medical, and legal domains. Additionally, the Guatemalan government recognizes the official orthography of Kaqchikel and all other Mayan languages in Guatemala as developed and endorsed by the *Academia de las Lenguas Mayas de Guatemala* 'Academy of the Mayan Languages of Guatemala' (ALMG) (Guatemala 1987). The data in this paper come from ongoing thesis study of the third language acquisition of Kaqchikel's sound system, specifically its stop consonants. All spoken data presented are produced by first language speakers of Kaqchikel who also speak Spanish and English as additional languages.

2.2 Kaqchikel phonology

2.2.1 Phoneme inventory

An understanding of the phoneme inventory is crucial to answering the question regarding the phonological insertion of a laryngeal feature. The phoneme inventory of Kaqchikel consists of up to 32 distinct phonemes. Of these 32, 22 are consonants, which are shown in Table 1.

Table 1: Kaqchikel consonant phonemes

	Labial	Alveolar	Palatal	Velar	Uvular	Glottal
Plain stop	p	t		k	q	ʔ
Glottalized stop	ɓ̥	t'		k'	ʠ̥	
Plain affricate		ts	tʃ			
Glottalized affricate		ts'	tʃ'			
Fricative		s	ʃ	x		
Nasal	m	n				
Lateral approximant		l				
Approximant		ɾ				
Glide			j	w		

The majority of the consonants of Kaqchikel are obstruents, as there is one series of fricatives at three places of articulation (alveolar /s/, palato-alveolar /ʃ/, and velar /x/), two pairs of affricates, and two series of stops at four places of articulation (bilabial, alveolar, velar, and uvular), as well as a phonemic glottal stop. There is no phonemic glottal fricative, indicative of the absence of contrastive [sg] in Kaqchikel.

Kaqchikel stops and affricates, like those of all other Mayan languages (Bennett 2016), are differentiated by the laryngeal feature [cg]. One series – what I'll call the plain stops – have a [−cg] feature. These are the stops /p t k q/ and the affricates /ts/ and /tʃ/. The other series, which I'll refer to as glottalized, is specified [+cg] and features mainly ejectives /t' k'/, but voiceless implosives do surface, particularly the bilabial /ɓ̥/ and uvular /ʛ̥/ (Patal Majzul et al. 2000). The glottalized affricates are both ejectives /ts' tʃ'/.

The six remaining sonorants can be grouped in a number of different ways, but most relevant to the current study is the distinction between nasals /m n/ on the one hand and non-nasal sonorants /l ɾ j w/ on the other.

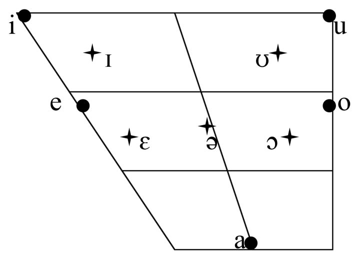


Figure 2: Kaqchikel vowel phonemes; Vowels indicated by a star are lax vowels not distinguished in all varieties of Kaqchikel

Kaqchikel minimally has a standard five-vowel system with high front /i/, mid front /e/, high back /u/, and mid back /o/ plus a single low vowel /a/. In addition to these five vowels are a series of five lax vowels, each of which has a tense counterpart among those five standard vowels. This tense-lax distinction developed from a short-long distinction in proto-K'ichee'an, one still exhibited in other K'ichee'an languages (Bennett 2016). However, the lax vowels of Kaqchikel are limited as to where within a word they may surface, namely only in stressed, word-final syllables (Rill 2013). If a lax vowel were to be dislocated outside of that position, it would neutralize to its tense counterpart. Furthermore, the distinction between each of the tense-lax pairs is not present in all varieties of Kaqchikel. Therefore, in stressed, word-final syllables, Kaqchikel has 5–10 tense and lax vowel phonemes, as shown in Figure 2, but only 5 tense vowels outside of those prominent contexts.

2.2.2 Words and syllables

Brown et al. (2006: 138) describe Kaqchikel as having word stress and that that word stress "is generally on the final syllable of a word". Bennett (2016), reporting on Mayan languages more broadly, states that "final stress is the norm in K'iche[e']an languages" (2016: 495) (example 1a). This is true even when the final syllable is wholly or partly a suffix (example 1b).

(1) a. *chiköp*
[tʃi.ˈkɔpʰ]
chiköp
animal
'animal'

b. *chikopi'*
[tʃi.ko.ˈpiʔ]
chiköp-i'
animal-PL
'animals'

In multi-word phrases, Brown et al. (2006) describe Kaqchikel as having phrasal prominence on the final syllable of the final word of every phrase. The realization of this prominence differs depending on the type of phrase, with declaratives and content (wh-)questions having a falling tone over the final syllable of the final word (Nelson 2020), while polar (yes/no) questions instead have a rising tone (Brown et al. 2006).

The prosodic emphases above show that Kaqchikel consistently places prominence on the right edge of prosodic domains. This becomes crucial when discussing allophony at the right edge in the following section.

3 Kaqchikel allophony

3.1 Stop allophony

In Kaqchikel, the right edge of words is also the conditioning environment for position-based allophony of several sound classes. In this subsection I present and discuss allophony of stops. Then, in the following section I show sonorants and their allophony.

As mentioned in §2.2.1, there are four plain stop consonants in Kaq-chikel. These are bilabial /p/, alveolar /t/, velar /k/, and uvular /q/. Plain stops may and do appear in either position of a syllable (onset or coda) and any position of a word (initial, medial, or final), and appear in roots as well as affixes. However, in word-final position, the allophone that surfaces is not a plain, voiceless, unaspirated stop such as [p] (Figure 3), but instead an aspirated stop [pʰ] (Figure 4).

Of course this is not restricted to just the bilabial stop /p/. It also occurs with alveolar /t/: [t] surfaces in non-final positions (Figure 5) while [tʰ] does so in word final position (Figure 6); the plain velar stop /k/: unaspirated or lightly aspirated [k] at the beginning of a word (Figure 7) vs. heavily aspirated [kʰ] at the end of a word (Figure 8); and the plain uvular stop /q/: lightly aspirated [q] in onset (Figure 9) and heavily aspirated [qʰ] in word-final coda (Figure 10).

Through these examples, we see that all plain stops in Kaqchikel exhibit positional allophony. In onset position, stops surface as voiceless and unaspirated (T). In word-final position, the same stop phonemes instead surface as voiceless and heavily aspirated (Tʰ).

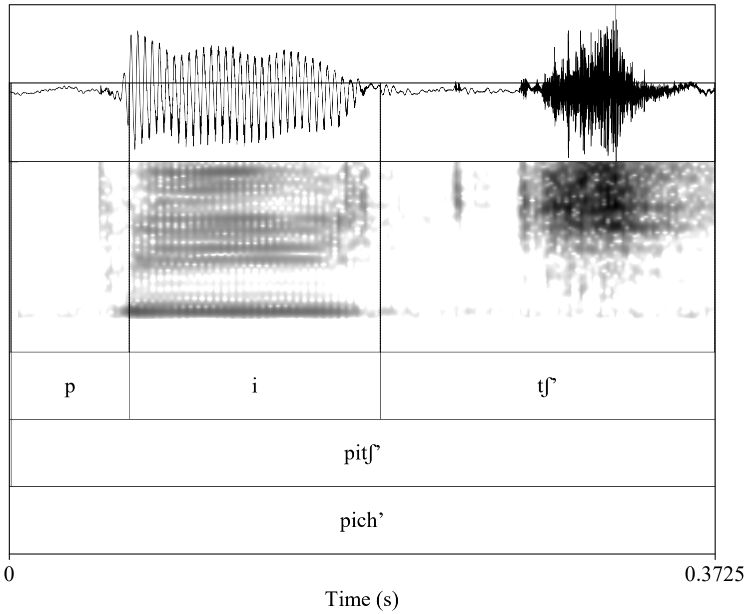


Figure 3: *pich’* /pitʃʼ/ [ˈpitʃʼ] ‘tender corn’ spoken by NKS1

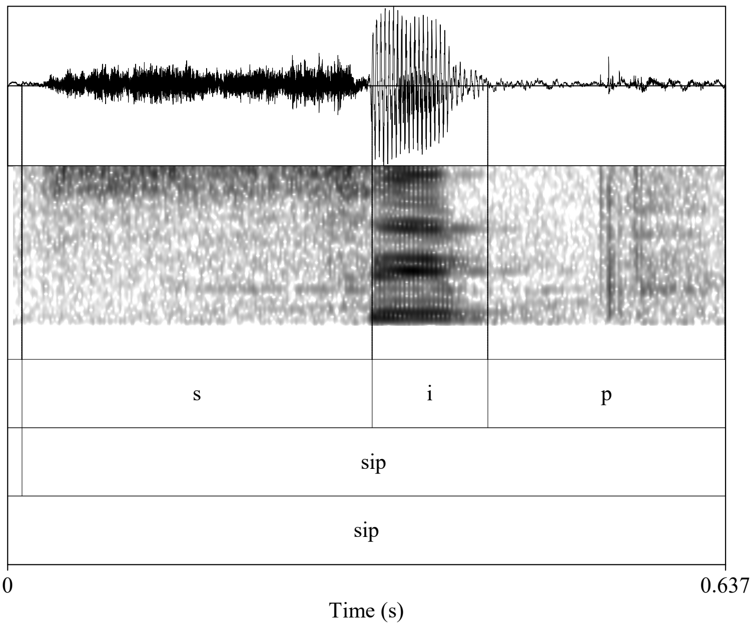


Figure 4: *sip* /sip/ [ˈsip^h] ‘tick’ spoken by NKS1

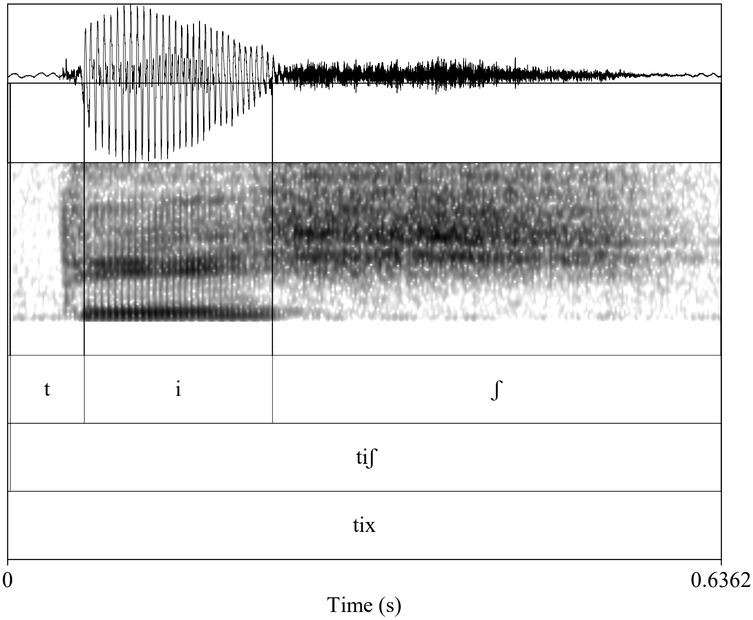


Figure 5: *tix* /tiʃ/ [ˈtiʃ] ‘elephant’ spoken by Kawoq

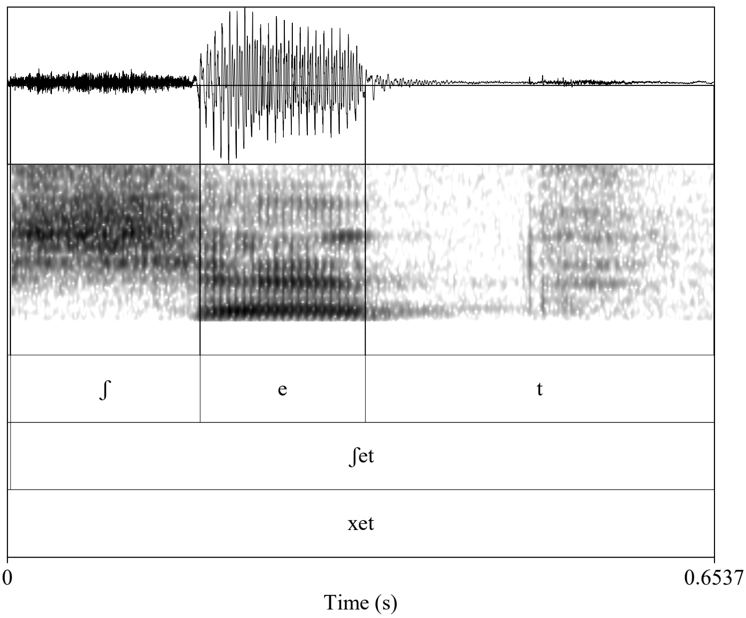


Figure 6: *xet* /ʃet/ [ˈʃetʰ] ‘hair whorl’ spoken by Kawoq

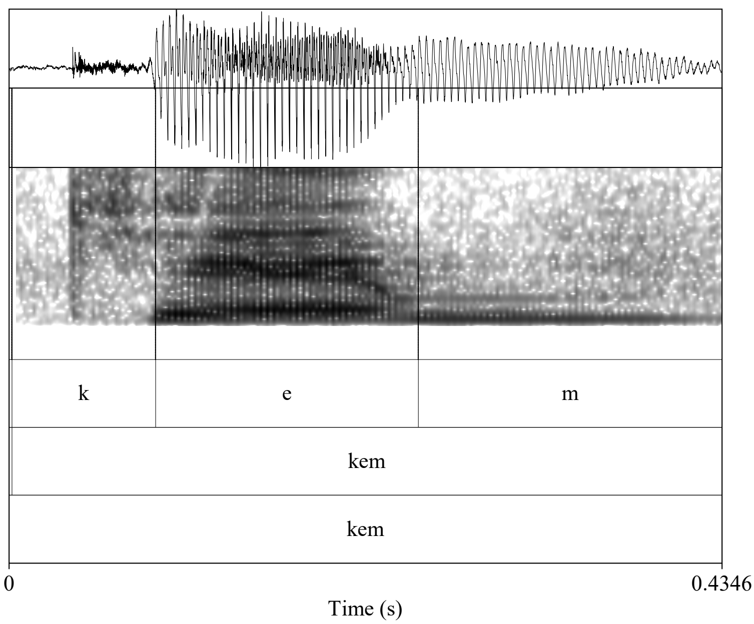


Figure 7: *kem* /kem/ [ˈkem] ‘weaving’ spoken by NKS1

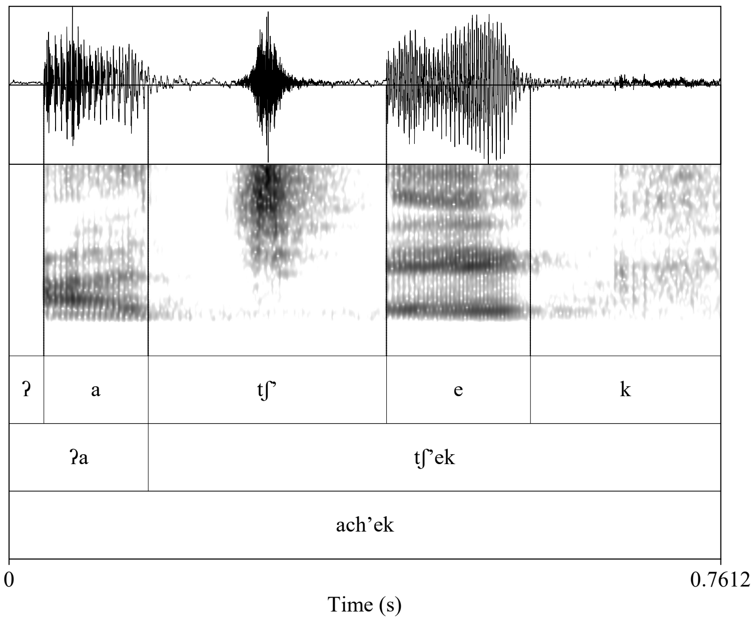


Figure 8: *ach’ek* /ʔatʃ’ek/ [ʔa.ˈtʃ’ek^h] ‘dream’ spoken by NKS1

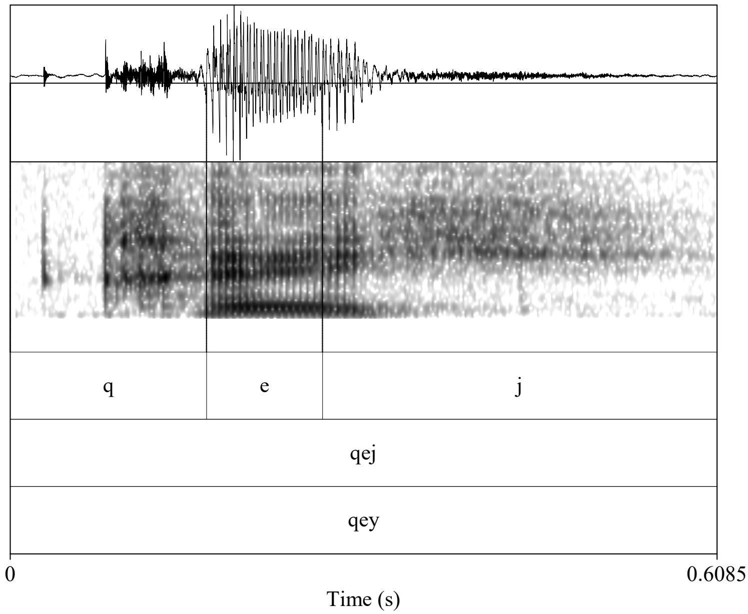


Figure 9: *qey* /qej/ [ˈqeç] 'our teeth' spoken by Kawoq

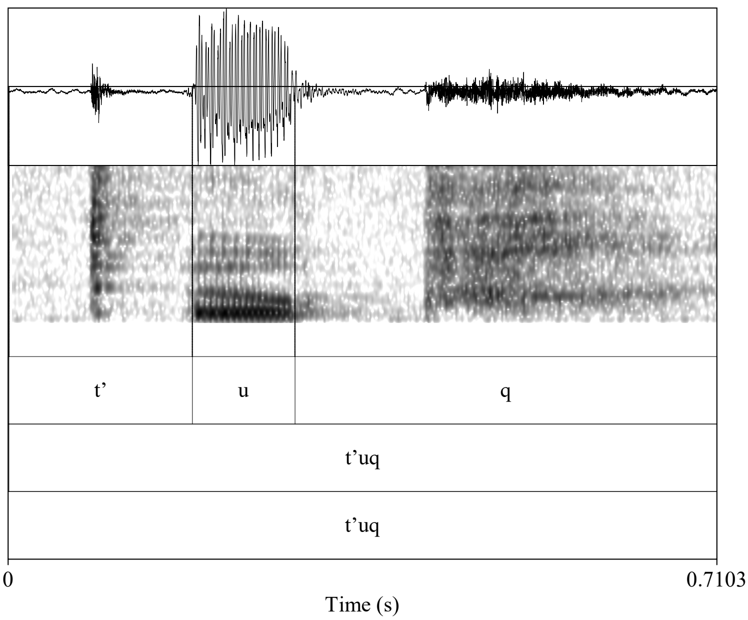


Figure 10: *t'uq* /t'uq/ [ˈt'uq^h] 'setting hen' spoken by Kawoq

3.2 Sonorant allophony

The four non-nasal sonorants (continuant sonorants) /l ɾ j w/ also exhibit positional allophony at the ends of words in Kaqchikel. However, it is realized differently than adding an aspiration burst to the end of the articulation of the voiceful sonorant, as is the case for the plain stops.

Examine the following examples, both of which contain the root *-söl* /-sɔl/ 'to tie'. The key segment here is the lateral /l/. In Figure 11, I show the third person plural antipassive *yesolon*, where /l/ surfaces as [l] in medial onset position. Compare this to Figure 12 *nkisöl*, where /l/, now in final position, surfaces as a voiceless alveolar lateral fricative [ɬ].

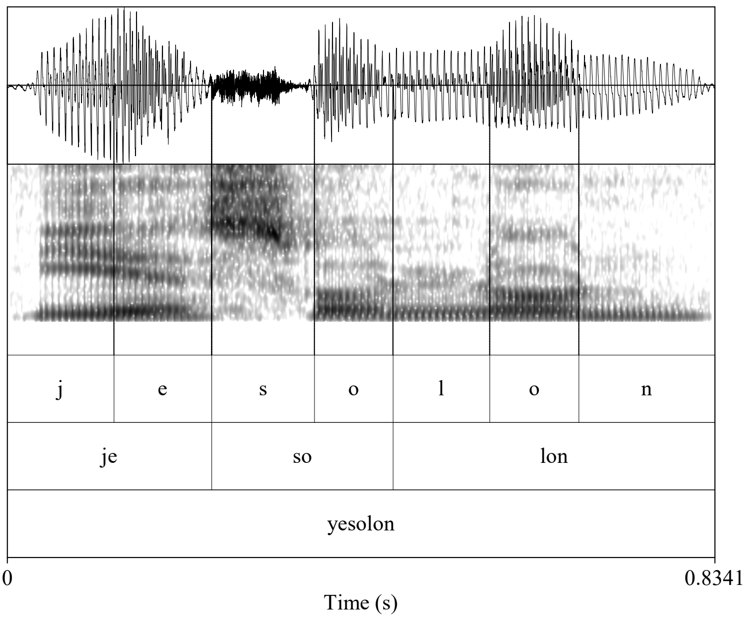


Figure 11: *yesolon* /jesolon/ [je.so.ˈlon] 'they untie' spoken by B'alam

Again, looking to other members of this class, similar patterns emerge. With the rhotic /ɾ/, a tap (optionally a trill) surfaces in onset position, as seen in Figure 13, but in word-final coda, this phoneme surfaces as a voiceless fricative [ʂ̩] (Figure 14).

Next, we have the glides /j/ and /w/. Figure 15 shows /j/ surfacing as [j] in onset, while Figure 16 shows this sonorant phoneme becoming a voiceless obstruent [ç] in word-final position. Similarly, /w/ is a (voiceful) sonorant glide as an onset (Figure 17), but a voiceless fricative as a word-final coda (Figure 18).

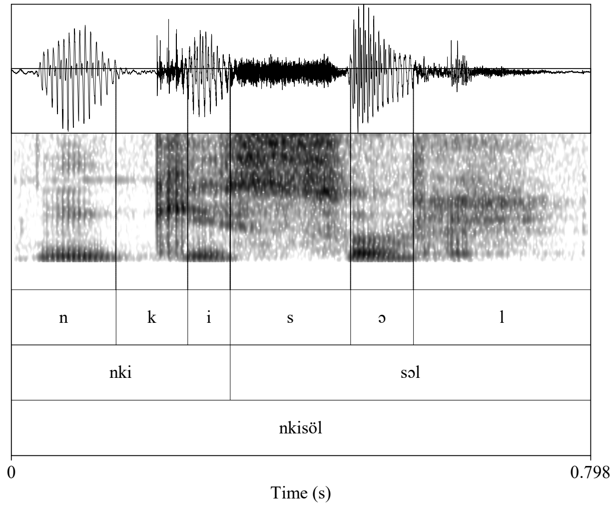


Figure 12: *nkisöl* /nkisɔl/ [nki.ˈsɔɫ] ‘they untie it’ spoken by B’alam

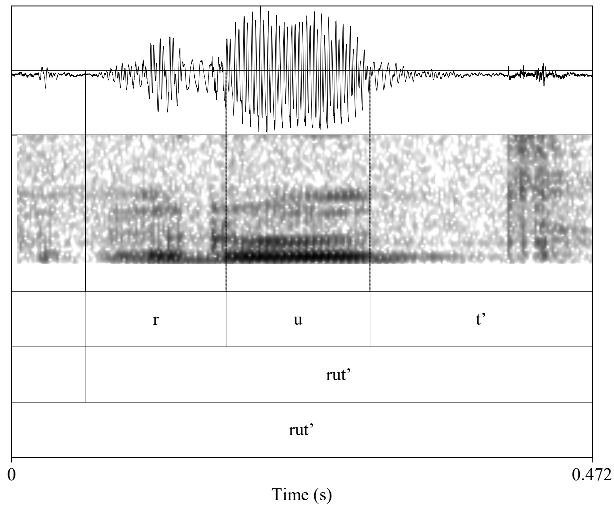


Figure 13: *rut’* /ɾut’/ [ˈɾut’] ‘receipt’ spoken by Yab’un

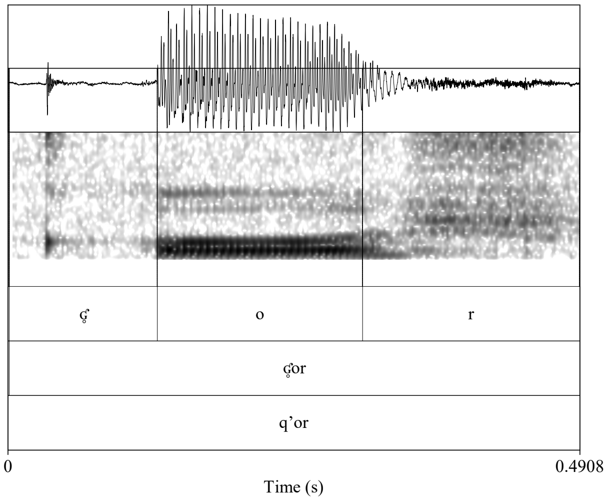


Figure 14: *q’or* /ʠ̥oɾ/ [ˈʠ̥oʂ] ‘atole (corn beverage)’ spoken by Yab’un

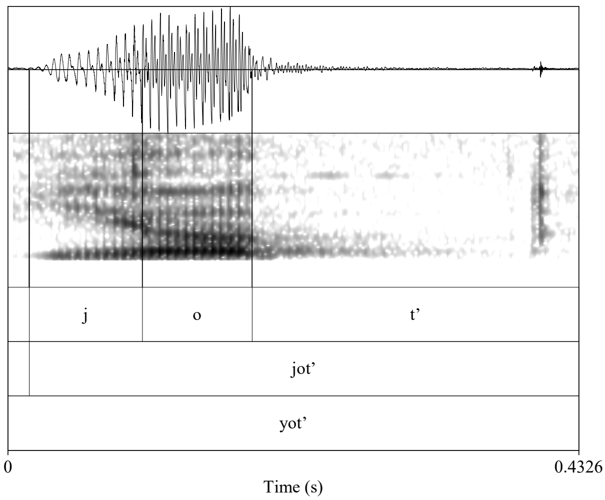


Figure 15: *yot’* /jot’/ [ˈjot’] ‘dimple’ spoken by B’alam

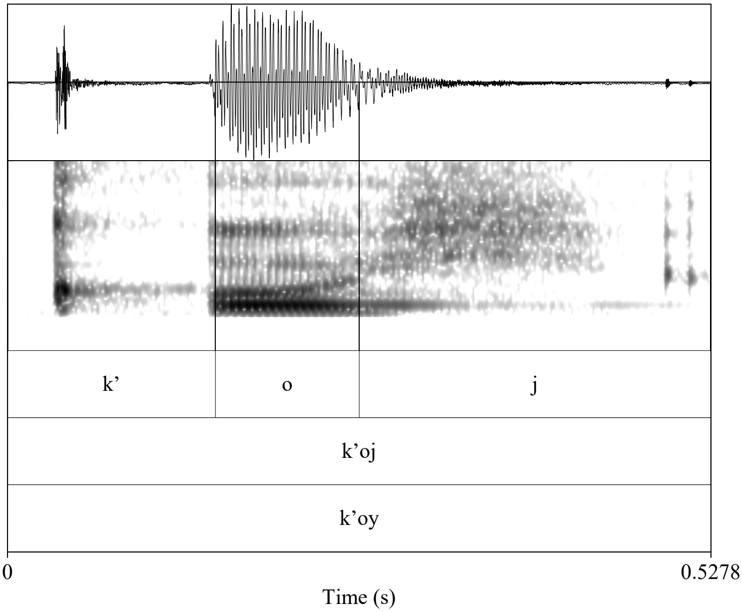


Figure 16: *k’oy* /k’oj/ [ˈk’oç] ‘spider monkey’ spoken by B’alam

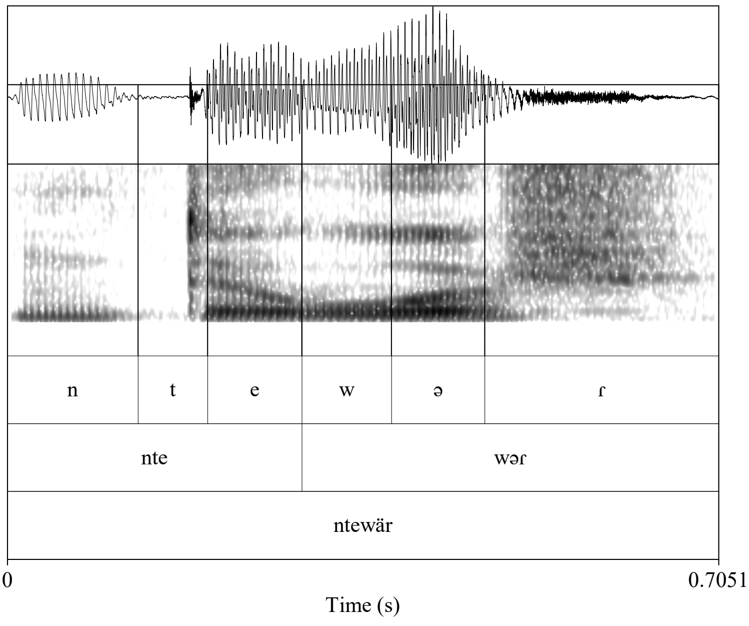


Figure 17: *tew* /tew/ [ˈtef] ‘cold’ spoken by Kawoq

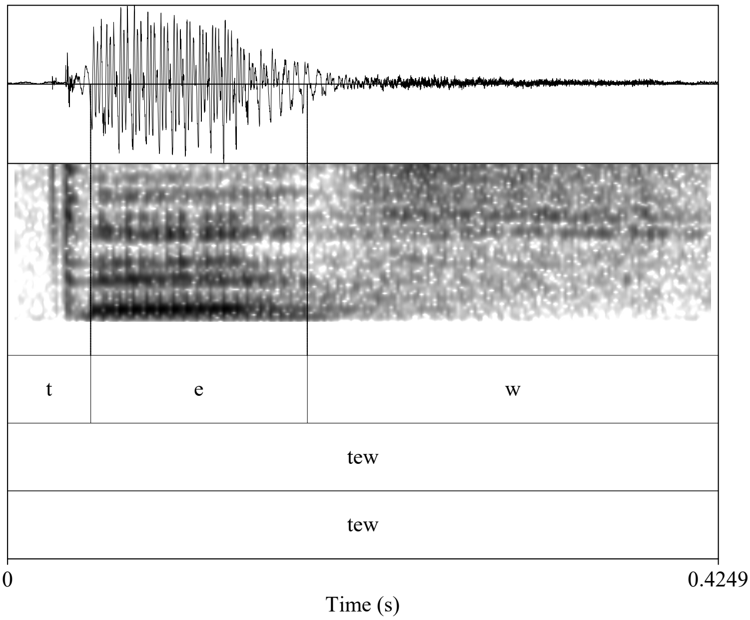


Figure 18: *ntewär* /ntewəɾ/ [nte.ˈwəʂ̬] ‘it gets cold’ spoken by Kawoq

3.3 Other sounds in final position

In the previous subsection, I showed that plain stops and non-nasal sonorants both exhibit particular cases of positional allophony where word-final occurrences differ substantially from non-word-final counterparts. What happens to the other sound classes in word final position? The answer to this question helps circumscribe the main pattern of allophony at the right edge of words in Kaqchikel. Therefore, in this subsection, I show examples of these sound classes in these positions.

First, let’s look at a glottalized stop. In Figure 19 the glottalized alveolar stop /t/ occurs in both (word-medial) onset and word-final coda. Both occurrences result in surface forms that are voiceless ejectives [t’]. Other glottalized stop phonemes behave similarly.

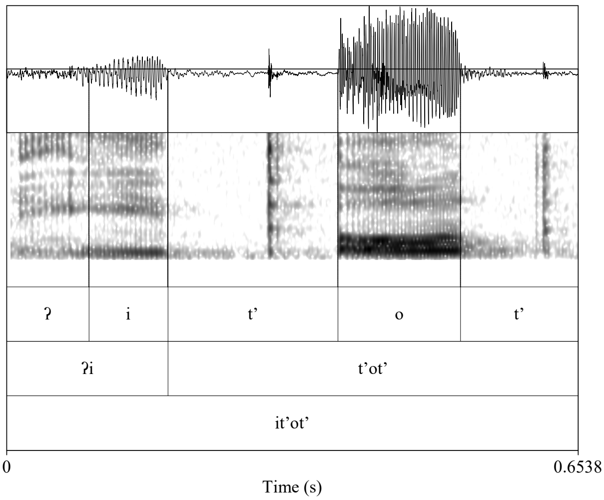


Figure 19: *it’ot’* /ʔit’ot’/ [ʔi.ˈt’ot’] ‘y’all’s conch’ spoken by NKS1

Next, consider the palato-alveolar fricative /ʃ/. It appears as an onset in Figure 20, but as a word-final coda in Figure 21. Again, both surface forms are the same: [ʃ], the voiceless palato-alveolar fricative. There is no positional allophony for fricatives, though word-final fricatives are considerably longer than fricatives in other word positions.

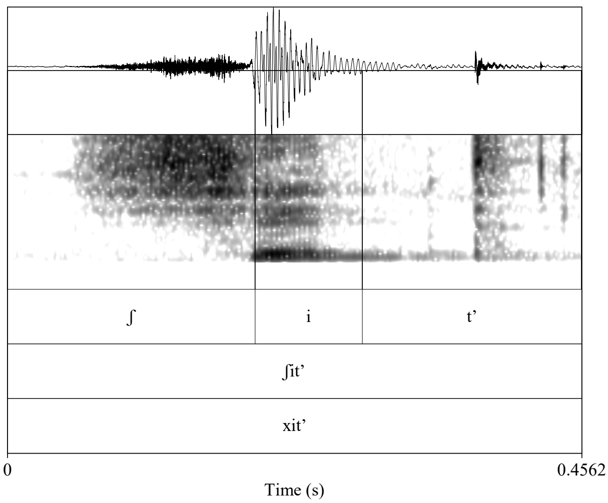


Figure 20: *xit’* /ʃit’/ [ˈʃit’] ‘It was filled well.’ spoken by Aq’ab’al

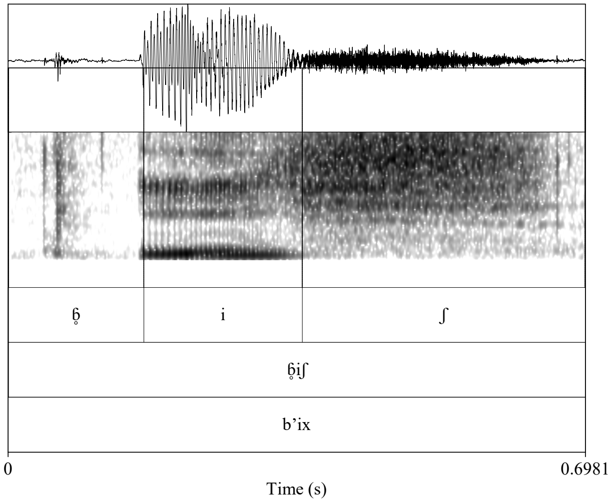


Figure 21: *b'ix* /ɓ̥iʃ/ [ˈɓ̥iʃ] 'song' spoken by Aq'ab'al

Finally, I show an example of the labial nasal /m/. Figure 22 has /m/ as it appears in onset, while Figure 23 has it in word-final coda. There is no difference between them other than the longer duration exhibited by the final [m] in Figure 23. The other nasal phoneme /n/ behaves similarly.

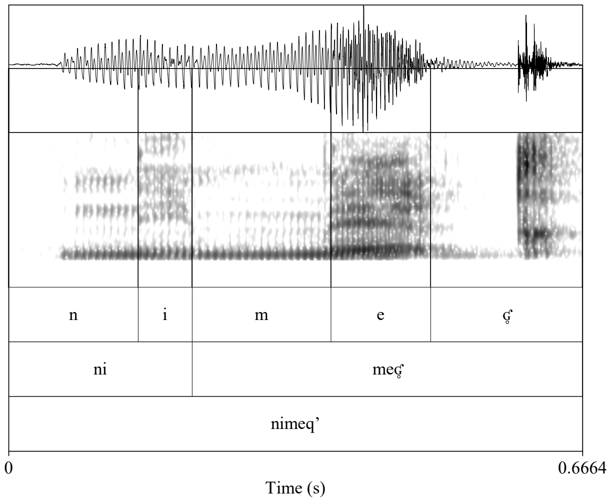


Figure 22: *nimeq'* /nimeʛ̥/ [ni.ˈmeʛ̥] 'It is warmed.' spoken by Aq'ab'al

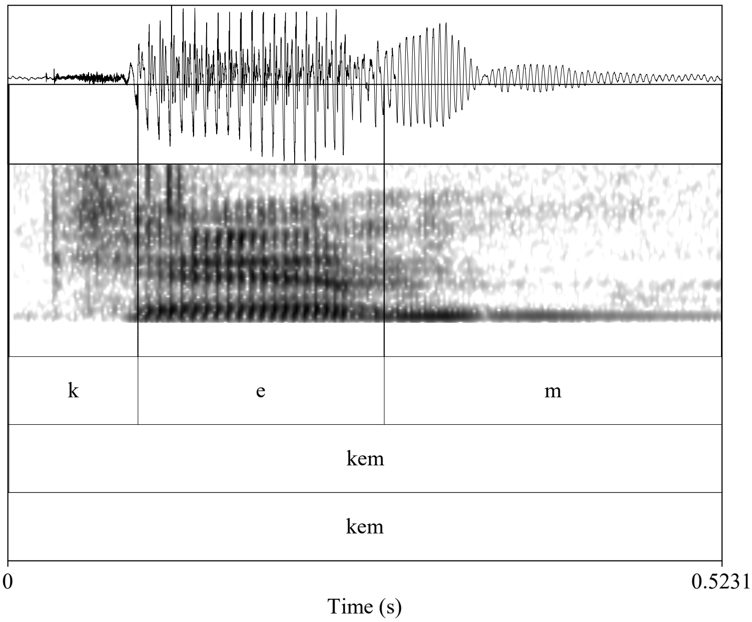


Figure 23: *kem* /kem/ [ˈkem] 'weaving' spoken by Aq'ab'al

Finally, there is the case of the alveolar and palato-alveolar affricates.[1] As with the previously discussed stops, these affricates do contrast on [cg], with the glottalized affricates surfacing as ejectives [ts' tʃ'] in both onset and word-final coda positions. Examples of the glottalized palato-alveolar affricate /tʃ'/ are shown in Figures 24 and 25.

The plain, non-glottalized affricates /ts tʃ/, on the other hand do exhibit positional variation, with word-final affricates being aspirated, aligning with descriptions of other Mayan languages such as Ch'ol (ctu) (Warkentin & Brend 1974) and Yukatek (yua) (AnderBois 2011) which have all plain oral stops and affricates aspirating in word-final position. However, as can be seen in Figures 26 and 27, the noisy period of frication and aspirtion after the release burst is longer in word-final positions. Thus, plain affricates, patterning phonologically with plain stops, do participate in a similar process to those stops, aspirating in word-final positions.

[1]As pointed out by an anonymous reviewer, affricates were notably missing from this analysis when I originally presented it at the 2021 Epenthesis Workshop. With many thanks to this reviewer, I have added the following data on glottalized and plain affricates to complete this survey of Kaqchikel's consonant inventory.

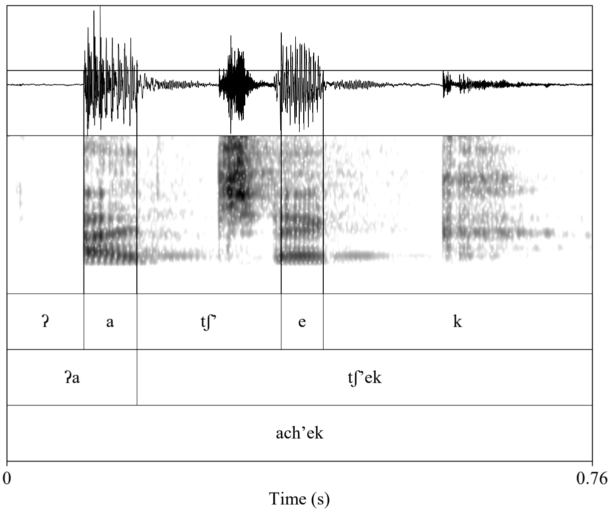


Figure 24: *ach'ek* /ʔatʃ'ek/ [ʔa.ˈtʃ'ekʰ] 'dream' spoken by B'alam

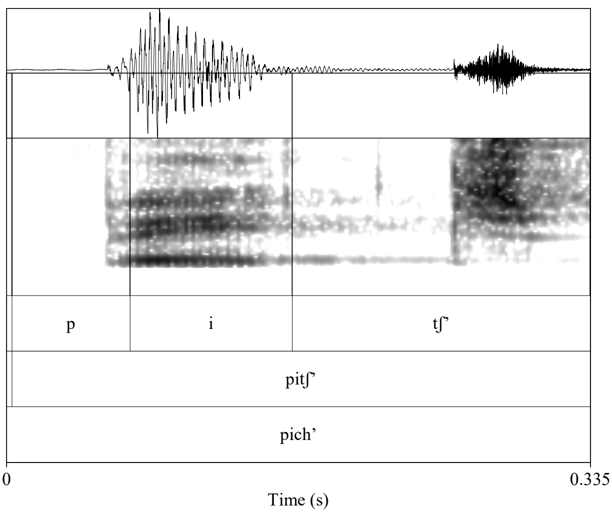


Figure 25: *pich'* /pitʃ'/ [ˈpitʃ'] 'tender corn' spoken by B'alam

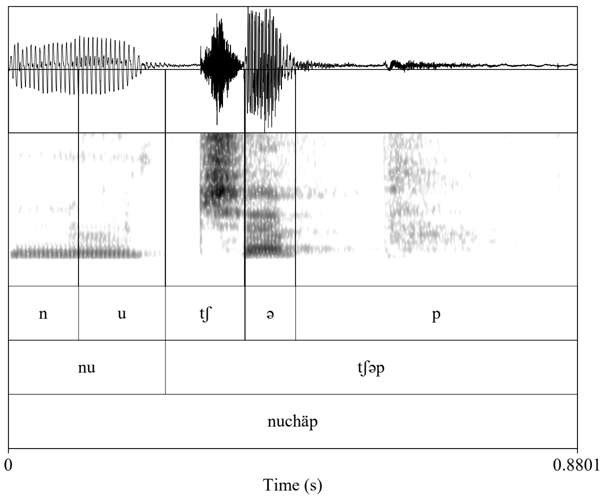


Figure 26: *nuchäp* /nutʃəp/ [nu.ˈtʃəpʰ] ‘I grab it.’ spoken by B’alam

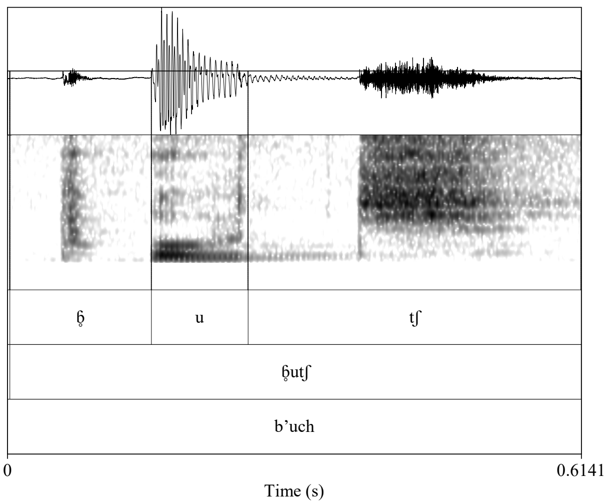


Figure 27: *b’uch* /ɓ̥utʃ/ [ˈbutʃʰ] ‘belly of an animal’ spoken by B’alam

The examples in this subsection have shown that despite the difference between the final and non-final allophones of the plain stop (including affricates) and sonorant classes, the other consonantal classes of the language do not exhibit similar types of allophony. The glottalized stops, (voiceless) fricatives, and (modally voiced) nasals are just that: glottalized stops, fricatives, and nasals in all positions. Thus, the positional allophony of Kaqchikel consonants is phonologically restricted to those first two classes: plain stops and non-nasal sonorants. In the next section, I offer a proposal as to why this is the case.

4 Proposal

4.1 [+spread glottis] insertion

To account for the facts that in Kaqchikel, certain sound classes (plain stops and sonorants), but not others (glottalized stops, fricatives, nasals) exhibit positional allophony at the right edge of words, I propose that this prosodic domain boundary is being marked (cf. Rubin & Kaplan 2024 [this volume] for discussion of prosodic domain specific epenthesis). Specifically, the right edge of every word in Kaqchikel is potentially marked by the insertion of a [+sg] feature.

Similar to languages like German (Iverson & Salmons 2007), [+sg] insertion in Kaqchikel causes unmarked stops to become aspirated. In that language, [+sg] causes final fortition of laryngeally unmarked stops [T] into aspirated stops [T^h]. In German, this neutralizes the distinction between them and the already [+sg] marked aspirated stops [T^h] present in the language's consonant inventory. In Kaqchikel, when [+sg] is inserted into a plain stop [T], it also creates an aspirated stop [T^h] (i.e. /p/ becomes [p^h], /t/ becomes [t^h], /k/ becomes [k^h], and /q/ becomes [q^h]). However, as the contrastive laryngeal feature in Kaqchikel is [cg], there is no contrast neutralization as there is in German. Instead, the contrast between plain stops [T] and glottalized stops [T'] is enhanced in word final position. These differences in positional allophony are shown in Table 2.

Table 2: Final [+sg] insertion comparison

<table>
<tr><th></th><th colspan="2">German</th><th colspan="2">Kaqchikel</th></tr>
<tr><th></th><th>Non-final</th><th>Final</th><th>Non-final</th><th>Final</th></tr>
<tr><td>Unmarked</td><td>T</td><td rowspan="2">T^h</td><td>T</td><td>T^h</td></tr>
<tr><td>Marked</td><td>T^h</td><td>T'</td><td>T'</td></tr>
</table>

For the Kaqchikel non-nasal sonorants, when they appear in the word-final position where they receive the [+sg] feature, they cannot so easily become aspirated stops. Instead, these continuant sounds become obstruents. Continuant obstruents are fricatives, and fricatives with [+sg] are voiceless fricatives (Vaux 1998). Thus each sonorant becomes a voiceless fricative at the same place of articulation: /l/ becomes [ɬ], /ɾ/ becomes [s̺], /j/ becomes [ç], and /w/ becomes [f].

This insertion of [+sg] fails to cause positional allophony of the other classes of consonants, even though they can surface in word-final positions. This can be explained for each of the consonant classes. First, the glottalized consonants already have an antithetical laryngeal specification of [+cg] making insertion of [+sg] impossible for those consonants. Next, the fricatives are already continuant obstruents. Insertion of [+sg] on these fricatives does not cause any change to the sounds, as voiceless fricatives are the end result of continuants being specified for [+sg] (Vaux 1998). Finally, the nasals are inherently voiced, despite [voice] not being a distinctive feature for consonants in Kaqchikel. This is due to licensing restrictions against laryngeal specification when [nasal] is specified (Itô et al. 1995). Thus [+nasal] sounds may not accept the insertion of [+sg] at the ends of words in Kaqchikel.

4.2 Implications

The proposal given in §4.1 carries several key implications for Kaqchikel phonology and phonological theory more generally. These implications arise from three points: the non-contrastiveness of [sg] in Kaqchikel, the selective insertion of [+sg] dependent upon consonantal class, and the enhancement of the laryngeal contrast between the plain and glottalized stops in word-final position. This subsection addresses each of these points in turn.

The main research question raised in §1 asks whether non-contrastive features can be active in the marking of prosodic domains. The data shown in the previous section demonstrate that [sg] is active in domain marking, despite not being active in distinguishing the phonemes of Kaqchikel. Unless this is a purely post-phonological, prosodic process, then this would challenge the tenet of the Contrastivist Hypothesis that only contrastive features are active during phonological computation (Hall 2007, Dresher 2009). I discuss alternative analyses that maintain the Contrastivist Hypothesis in §5.2.

The possibility that this is a post-phonological, prosodic process is complicated by the fact that [+sg] insertion is phonologically selective. Only some of the consonantal classes readily accept the new laryngeal feature, namely the plain (laryngeally unspecified) stops and non-nasal (continuant) sonorants. The other consonantal classes do not accept [+sg]; the glottalized stops are already specified for an opposing [+cg] laryngeal feature, while the nasals do not license [+sg], but instead [+voice] in all positions. This shows that the insertion process is more phonologically controlled than it would be under a simple post-phonological, prosodic process.

Finally, the Kaqchikel data illustrate a process of contrast enhancement, differing from the case of German contrast neutralization. Enhancement, after Keyser & Stevens (2006), refers to cases where distinctive features, which are always accompanied by phonetic gestures, receive secondary gestures in order to supplement its insufficiently salient primary gestures. In this case, the aspiration gestures of [+sg] are introduced to plain stops in final position, leading to enhancement of the contrast between those plain stops and the laryngeally marked glottalized stops in that position. Whether this enhancement was the initial impetus for the insertion of [+sg], that subsequently spread to the sonorant class, is a matter for future research.

5 Discussion

5.1 Alternative explanation: Element Theory

5.1.1 Introduction to Element Theory

The insertion of [+sg] is not the only possible explanation for the positional allophony observed in Kaqchikel. Indeed, Nasukawa et al. (2018) examine Kaqchikel positional allophony of stops and sonorants under the framework of Element Theory. In this subsection I discuss Element Theory and their proposal so that I can then compare it to the one I offer in §4.

Harris & Lindsey (1995) present Element Theory as a theoretical framework of phonology that represents phonological contrasts by way of elements rather than features. Instead of more than a dozen *SPE*-like features, Element Theory uses six elements to represent all of the contrasts of spoken languages. To accomplish this feat, instead of a feature having a relatively restricted set of accompanying gestures and cues, multiple instances of an element may combine within a segment to call for the necessary cues of the language (Nasukawa et al. 2018). Therefore, the elements are founded upon the acoustic properties of sounds.

Each of the six elements of Element Theory is typically represented as single uppercase letter (e.g. |A|) and can be bundled together with other elements using brackets (e.g. [AH], Harris & Lindsey 1995. Each element is associated with one or more consonant categories as well as one or more vowel categories.

Three of the elements are founded upon properties of resonance (Nasukawa et al. 2018). These are |A|, |I|, and |U|. The element |A| is labeled by Nasukawa et al. (2018) as “Mass” for its mass of spectral energy and is associated with uvular or pharyngeal consonants and non-high vowels. |I| is called “Dip” for its spectral trough and is associated with dental and palatal consonants as well as front vowels. The final resonance element is |U| “Rump”, which is associated with both labial and velar consonants plus rounded vowels. These elements together perform roles similar to place features in featural theories of phonology.

The remaining three elements, |ʔ|, |H|, and |N| , are for non-resonant properties and are of primary importance to the current issue of Kaqchikel allophony (Nasukawa et al. 2018). |ʔ| carries the label “Edge” and provides the occlusion property for obstruents, as well as laryngealization of creaky voice vowels. |H| is called “Noise” and is associated with voicelessness, aspiration, and frication in consonants, but high tone in vowels. Finally |N| “Murmur” refers to nasality and voicing of consonants, and when it is part of a vowel it may also signal nasality or low tone. These elements act in ways that manner and laryngeal features do in featural theories.

5.1.2 Element Theory in Kaqchikel

Based on these descriptions of the elements alone, one can see that |ʔ| and |H| are going to be most relevant for Kaqchikel allophony. |ʔ| is indeed the difference between the plain and glottalized stops of Kaqchikel, with the glottalized series having two (one for occlusion and one for glottalization) and the plain series having one (for occlusion), Additionally all voiceless obstruents also carry |H|.

Nasukawa et al. 2018 propose, however, that in addition to these base attributions, the elements are also at work in creating the positional allophony observed in Kaqchikel. Specifically, they claim that the right edge of the domain is prominent in Kaqchikel, similar to how the left edge of prosodic domains is prominent in English. Thus, similar to how English voiceless/lenis stops aspirate at the left edge of words, Kaqchikel plain stops aspirate at the right edge of words. Therefore, they propose, a fortition element of |H| is inserted in the prominent, domain-final position.

It should be noted, as Nasukawa et al. (2018: 629) do: “not all versions of Element Theory allow multiple tokens of an element in an expression”, however the version Nasukawa et al. (2018) present for their analysis of Kaqchikel (named Precedence-free Phonology) does. While other models of Element Theory, including Harris & Lindsey (1995), mark headedness on elements (visualized as underlining of the element) in order to account for finer distinctions in acoustic properties among segments, Nasukawa et al. (2018) accomplish this distinction through the use of multiple tokens of an element. In the case of unaspirated voiceless stops versus aspirated voiceless stops, the former has a single |H| while the latter has two |H|, corresponding to the extra noise component present in stop aspiration.

The insertion of the fortition element |H|, by its nature within Element Theory, causes various outcomes. The clearest of these is aspiration of stops. Moreover, Nasukawa et al. (2018) propose, |H| causes the change of voiceful sonorants into voiceless fricatives. However, one |H| element is not enough to cause this change, as the sonorants do not carry a |H| on their own. One |H| element would lead to voiceful/lenis fricatives. But voiceful fricatives are not produced; voiceless ones are. As Nasukawa et al. (2018: 630) propose: "voiceless fricatives and aspirated stops are represented … each with two tokens of |H|". Therefore, an additional |H| element would be required to form the observed voiceless (fortis) fricatives.

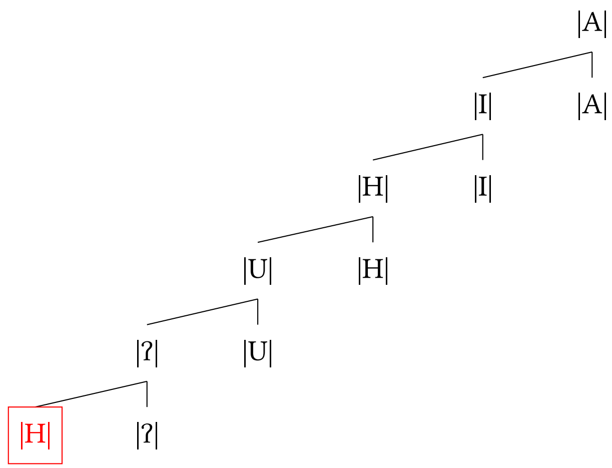


Figure 28: Change in structure of [ik] to [ikʰ] via boundary marking, cf Figure 6 in Nasukawa et al. (2018). The added element |H| (shown and boxed in red) indicates aspiration of the stop.

Through this fortition process, |H| is involved in the domain-final allophony of both stops and sonorants, but in different quantities. One fortition element |H| is added to turn the plain, voiceless stops into aspirated stops, but one structural element |H| is needed to turn the sonorants into fricatives and then another fortition element |H| is added to form voiceless fricatives. This contradicts the proposal of Nasukawa et al. (2018) that the addition of a *single* fortition element |H| domain-finally can account for allophony in both plain voiceless stops and sonorants in Kaqchikel. Furthermore, the tree structure diagram for [ik^h], shown here as Figure 28, shows *three* |H| element nodes: one for obstruency, one for voicelessness (doubled together in the tree), and one for aspiration. The linear representation they give for the consonantal portion of this syllable (i.e. [k^h]) is $[[[\text{Hʔ}]_{\text{ʔ}}\text{U}]_{\text{U}}\text{H}]_{\text{H}}$ (2018: 634). However, they give neither a tree structure nor a linear representation for the structure of any of the spirantized sonorants.

Expanding outside of the observed allophonic patterns in Kaqchikel, the analysis of Nasukawa et al. (2018) does not attempt to account for the fact that the other Kaqchikel consonant classes *do not* participate in similar cases of word-final allophony. If the Noise element |H| were introduced to any segment that appears in the relevant position (word-finally), it should have no problem combining with the structural elements already present.

For the glottalized stops, the structural element bundle is minimally [ʔ ʔ H]: one |ʔ| for the occlusion, another |ʔ| for the glottalization, and one |H| for their voicelessness.[2] The fortition element should attach to the structural elements of the glottalized stops. However, no allophony or change in elements is observed; Kaqchikel glottalized stops are still glottalized and unaspirated in final position. Therefore the fortition element |H| either has no effect on glottalized stops or is somehow prevented from being inserted when these sounds occur in the prominent position. However, there is no *a priori* reason Element Theory would prohibit this combination. Moreover, the structural elements making up voiceless glottalized stops already consist of this very combination.[3]

Phonemic voiceless fricatives appearing at the right edge should, however, have no issues combining their structural |H H| with the fortition element |H| from the domain boundary. Nevertheless, Nasukawa et al. (2018) make no mention of how the phonemic fricatives actually play into their analysis.[4]

Finally, we have the case of the nasals, the two of which do not exhibit allophony at the right edge.[5] Again, no discussion of them is given in Nasukawa et al. (2018), but presumably, the nasality of these sounds is represented by the |N| element. The possible effects of fortition element |H| on structural element

[2]The only mention of the structure of glottalized stops appears in a footnote to show that doubling of elements is not only permitted, but necessary to build up the phonemic contrasts exhibited by Kaqchikel (Nasukawa et al. 2018: 630).

[3]Rather it is |A| that Nasukawa et al. (2018) identify as incompatible with |H| due to incompatibility of their articulatory gestures.

[4]Nasukawa et al. (2018) describe Kaqchikel as having five contrastive fricative phonemes /s ʃ x χ h/, yet there are only three: /s ʃ x/ (cf Table 2.2.1 and Brown et al. 2006).

[5]Nasukawa et al. (2018) also describe Kaqchikel as having three contrastive nasals, adding /ŋ/ to the two /m n/ that I show in Table 1. This is not true of Kaqchikel. Alveolar /n/ *is* produced as [ŋ] in some environments, by some speakers, but not contrastively.

|N| are neither explained nor predicted by Nasukawa et al. (2018), and I struggle to provide an explanation of my own. I would expect, however, that devoiced/ aspirated nasals are a possible result of |N| + |H|. Devoiced nasals, however, do not occur here or in any position in Kaqchikel. So again, the domain boundary fortition element |H| appears not to have any effect on the class of nasal consonants in Kaqchikel.

To summarize, Nasukawa et al. (2018) propose that Kaqchikel plain stops and sonorants become their word-final allophones of aspirated stops and voiceless fricatives because the right-edge is prominent in Kaqchikel. This prominence is realized through the addition of a fortition element |H|. However, in order for this analysis to work, one |H| must be added to the voiceless stops to induce aspiration, while two |H| are added to the sonorants to induce obstruency and then voicelessness. This lack of parsimony, combined with the absence of explanation for the other consonants in Kaqchikel not exhibiting any similar patterns of allophony, makes the Precedence-free Phonology analysis of Nasukawa et al. (2018) less preferable than the current analysis that [+sg] insertion adequately derives the surface forms of all Kaqchikel consonant classes. This does not, however, preclude the possibility that another version of Element Theory can provide an adequate explanation for the observations of Kaqchikel.

5.2 Alternative explanation: Contrastivist Hypothesis

Further alternatives to the analysis presented here may support the Contrastivist Hypothesis of Dresher (2009).

As mentioned above, the Contrastivist Hypothesis holds that only contrastive features may participate in the lexical phonological processes of a language. "Lexical phonology," as Dresher (2009: 118) explains, "interacts with the morphology and the lexicon, and exhibits properties such as cylcic application, restriction to derived environments, and exceptions". In order for [sg] to participate in lexical phonology, it must be one of the contrastive features of the language's phonology. I do not take this to be the case in Kaqchikel.

However, as contrastive [sg] can be achieved within the Contrastivist Hypothesis. The Successive Division Algorithm (SDA) may produce feature sets not immediately apparent upon first glance at a language's phoneme inventory. If [sg] is used in dividing the contrastive segments of Kaqchikel early in the algorithm, namely in separating the three phonemic fricatives /s ʃ x/ from the rest of the [−sg] consonants by specifying the fricatives as [+sg]. Wax Cavallaro (2020) points this out as a possible solution for a similar case in Tz'utujil (described in greater detail in §5.3.1).

If this were the SDA path used in Kaqchikel, [sg] would indeed be a contrastive feature of Kaqchikel, and would therefore be able to participate within the lexical phonology of the language. However, separate positive evidence for the presence of contrastive [sg], other than the processes described in the current research, would be desirable before continuing with this analysis. For instance, if [sg] were a contrastive feature, then the lack of a phonemic glottal fricative /h/, as the placeless [+sg] consonant, would be curious (cf. the distribution of /h/ and aspiration in English and Korean as evidence for contrastive [sg] in those languages (Davis & Cho 2003)).

The basal specification of [+sg] for Kaqchikel fricatives additionally contradicts the proposal of Kehrein & Golston (2004: 2) that "[a]n onset, nucleus, or coda has a single, unordered set of laryngeal features". If both this proposal and the [sg] division of Kaqchikel fricatives were true, then clusters containing a fricative and a non-fricative would not be possible in Kaqchikel. However, these clusters are allowed to surface in both native words (example 2a) and loanwords (example 2b). Thus, I conclude, the proposed division of fricatives using [+sg] as their contrastive feature is untenable for Kaqchikel.

(2) a. ***xk’****is*
[ˈʃk’is]
ʃ-∅-k’is
‘it was finished’

b. *wa****kx***
[ˈwakx]
wakx
‘cow’

Another solution, which avoids the issue of [sg] being a contrastive feature of Kaqchikel altogether, is to place the processes of aspiration and spirantization outlined in §4, into the postlexical phonology of Kaqchikel. As further explained in Dresher (2009: 118): "Postlexical phonology follows the lexical phonology and does not observe the above restrictions having rather properties one would associate with phonetics rules". This could be the case in Kaqchikel, as contrast enhancement is often a phonetic, rather than phonological, process. However, as explained in §4.2, there are exceptions to the aspiration and spirantization rules of Kaqchikel. Exceptions are indicative of lexical phonology as opposed to the exceptionless rules of postlexical phonology. Indeed, the processes described in the current research depend on the phonological structure of the segments at the right edge of the word in Kaqchikel, placing them firmly within the scope of lexical phonology.

5.3 Other languages

5.3.1 Tz’utujil

Processes similar, yet not identical, to Kaqchikel’s allophony can be observed in related languages. Bennett (2016) reports that word-final plain stops are typically aspirated across the Mayan family, claiming it’s a regular, obligatory process for all but two Mayan languages. Additionally, final devoicing/spirantization of sonorants is found across the K’ichee’an, Tseltalan, and Huastecan branches of the family (Bennett 2016).

Wax Cavallaro (2020) analyzes a specific case very similar to the Kaqchikel case in Tz’utujil (tzj), a closely related K’ichee’an language, citing data from Dayley (1985). Of its identical consonant inventory, she also observes that plain stops aspirate in word-final position (example 3a), but also in word-internal codas (i.e. syllable-final position) (example 3b).

(3) a. Tz’utujil (Dayley 1985)
Word-final aspiration
tut
[tʊt^{h}]
‘palmera’

b. Tz’utujil (Dayley 1985)
Syllable-final aspiration
saqb’ach
[saq^{h}ɓat͡ʃ^{h}]
‘hailstone’

Similarly, Tz’utujil also has word-final spirantization/devoicing (example 4a) but also coda spirantization/devoicing more generally (example 4b). Thus, seemingly, the insertion of [+sg] is occurring at the right edge of every syllable in Tz’utujil, as opposed to the process only occurring at the right edge of every word (or every word-final syllable) in Kaqchikel.

(4) a. Tz’utujil (Dayley 1985)
Word-final devoicing
way
[waj̥]
‘tortilla’

b. Tz’utujil (Dayley 1985)
Syllable-final devoicing
Moysees
[moj̥seːs]
‘Moses’

A further point of departure from the Kaqchikel data, Wax Cavallaro (2020) also notes word-final devoicing is not restricted to the non-nasal sonorants: nasals also *partially* devoice in that position (example 5a). However, the general coda devoicing rule does not apply to nasals, as medial codas do not (example 5b).

(5) a. Tz'utujil (Dayley 1985)
Nasal devoicing
naan
[naːn͡n̥]
'lady'

b. Tz'utujil (Dayley 1985)
Nasal non-devoicing
xinwa'i
[ʃinwaʔi]
'I ate'

Wax Cavallaro (2020) provides a constraint ranking to derive all of these observations within an Optimality Theory framework. I do not discuss that part of her analysis here. More relevant to – and in agreement with – the current analysis of Kaqchikel is that [sg] is being manipulated by Tz'utujil phonology despite not being a contrastive feature in the language. However, the process of [+sg] insertion in Tz'utujil is less restricted/more general than that of Kaqchikel: [+sg] is inserted at all codas, not just word-final ones, though constraints exist that block its insertion on word-medial nasals.

5.3.2 Blackfoot

In addition to the similar cases of word-final allophony in related Mayan languages, there are cases of final aspiration/devoicing/spirantization across the world, including many other Indigenous languages of North America, including Plains Cree (Flynn et al. 2019), Cherokee (chr) (Montgomery-Anderson 2008), Nahuatl (nah) (Launey 1992), and Mistanlta Totonac (tlc) (MacKay 1994).[6]

In Blackfoot (bla), an Algonquian language spoken in the northwestern prairies of North America, Windsor (2016) analyzes a case of [+sg] insertion affecting phrase-final vowels, rather than word- or syllable-final consonants observed in Kaqchikel and Tz'utujil.[7] In Blackfoot, vowels devoice at the ends of orthographic words, which map to phonological words prosodically (example 6).

(6) Blackfoot vowel devoicing (Windsor 2016: 64 (3))
Ánniksi ákaímahkihkinaiksi inókiwa
ann-iksi áka-íímahkihkinaa-iksi ino-okiwa
[ánːiksi̥ ákɛ́ːmaxkiçkinɛksi̥ inókiʌʍḁ]
'those old sheep see us'

Moreover, Windsor (2016) explains, stops in Blackfoot are actually aspirated in the same phrase-final positions, though these instances are less frequent due to the relatively few consonant-final suffixes in the language (example 7).

[6] See Campbell et al. (1986) for discussion of final devoicing as a potential areal feature of Meso-America

[7] The fact that [+sg] insertion is not observed to affect Kaqchikel vowels is due to morphological constraints that all lexical words and suffixes end in a consonant.

(7) Blackfoot stop aspiration (Windsor 2016: 68 (based on spectrogram in 7b))
Piit!
pii-t
[pi:t^h]
'enter!'

In previous work, Windsor found that the average duration of the devoiced period in vowels and aspirated period in stops were roughly equivalent (Windsor & Cobler 2013). This leads to the notion that these two effects at the same phrase-final position are in fact part of the same process: one of [+sg] insertion.

Windsor (2016), in developing this analysis, proposes that [sg] is in fact contrastive in Blackfoot, despite the language's lack of contrastive aspiration or /h/ (or for that matter, voicing or glottalization indicative of other laryngeal features; cf. Elfner 2006). This claim is based on an reinterpretation of sequences previously described as velar fricative /x/ + stop /T/ as pre-aspiration.[8] This allows Windsor's proposal to satisfy the requirement of the Contrastivist Hypothesis that only contrastive features be active in phonology (Hall 2007, Dresher 2009).[9]

From there, Windsor (2016) proposes that [+sg] is inserted at the right edge of phonological phrases (orthographic words), leading to the observed allophony of voiced vs. devoiced vowels and plain vs. aspirated stops in phrase-final position.

5.3.3 Generalizing [spread glottis] insertion

The two additional cases of [+sg] insertion presented in the previous two subsections show that [+sg] insertion is found at the right edge of prosodic domains other than the word-final pattern observed in Kaqchikel. In Tz'utujil, [+sg] is inserted at the edge of every syllable, while in Blackfoot, it is inserted only at the edge of phonological phrases.

Iverson & Salmons (2007), building upon work in Evolutionary Phonology by Blevins (2004, 2006), show that marking of phrase boundaries with laryngeal cues (e.g. aspiration or glottalization) may spread to smaller domains through a process of generalization. So a process specific to the right (or left) edge of a phonological phrase can be generalized first to the respective edge of words within that phrase, and then to the edge of syllables within the word.

[8]Blackfoot stops contrast three ways for place: labial /p/, coronal /t/, and dorsal /k/ (Windsor 2016).

[9]This step would end up being an unnecessary step in the analysis if the current proposal regarding Kaqchikel and Wax Cavallaro's (2020) proposal regarding Tz'utujil are borne out.

The final proposal of the current study is that there exists a typology of [+sg] insertion processes at play in various languages' phonologies. Specifically, [+sg] may be inserted at the edge of various prosodic domains (e.g. the phrase in Blackfoot, the word in Kaqchikel, and the syllable in Tz'utujil.) This aspect of the typology comes about due to generalization of [+sg] insertion from the highest level (phrase) down to lower levels within the prosodic hierarchy. Additionally, this typology can be built upon which segments within the phonological inventory of the language are affected by the insertion of [+sg]. In other words, for some languages, only some classes of consonants are affected, while in other languages, vowels may be affected by [+sg] insertion. Again we see that this may develop/evolve over time: only plain stops and non-nasal sonorants are affected in Kaqchikel, while in Tz'utujil, nasal sonorants have begun to show phrase-final allophony as well.

6 Conclusion

The pattern of Kaqchikel allophony discussed in this paper, along with a similar pattern in Tz'utujil, show that the languages consistently mark the right edge of prosodic domains with laryngeal cues. These cues may be explained by the insertion of a laryngeal feature [+spread glottis]. This featural insertion has several implications for phonological theory.

The main implication of this proposal is that [+sg] is inserted despite [sg] not being a contrastive feature in the language. This conflicts with the Contrastivist Hypothesis of Hall (2007) and Dresher (2009). Under the current proposal, non-contrastive laryngeal features are available to be manipulated, or in this case inserted, by the phonology, with variable effects depending on the features already present in the segment receiving the feature.

Secondly, following Iverson & Salmons (2007), edge marking is predicted to spread and generalize in a single direction: from higher prosodic domains down to lower prosodic domains. Thus, I expect that historically, Kaqchikel only exhibited [+sg] insertion at the right edge of phrases, similar to the pattern seen in Blackfoot. This has since spread so that the insertion now occurs at the end of every word as well. In the future, this may generalize further to occur in every syllable coda, as it does in Tz'utujil.

Finally, this study surveyed three different patterns of [+sg] insertion to show that patterns can differ in various ways, including which prosodic domain the featural insertion applies to and which sounds are affected by the insertion. Future work on this topic will explore other cases of [spread glottis] or even [constricted glottis] insertion in order to explore and expand the breadth of the typology of laryngeal feature insertion.

Abbreviations

ALMG	Academia de Lenguas Maya de Guatemala
[cg]	[constricted glottis]
[sg]	[spread glottis]
SDA	Successive Division Algorithm
SPE	*The Sound Pattern of English* (Chomsky & Halle 1968)
UNESCO	UN Educational, Scientific and Cultural Organization

Acknowledgements

Personal acknowledgements

I offer special thanks to my Kaqchikel friends and teachers, Ixchel, Kawoq, Aq'ab'al, B'alam, Yab'un, and an anonymous contributor, among many others, whose productions I could analyze. This study is impossible without their devotion to their language and culture. I am also grateful to the attendees of various conferences and workshops where I've presented parts of this work. I specifically thank the attendees of the 2020 Meeting of the Society for the Study of the Indigenous Languages of the Americas for their thoughts on and references to similar patterns observed in other languages. I thank the attendees of the 2021 Epenthesis and Beyond workshop at Stony Brook University for discussion on possible causes for the observed allophony. Additionally, I thank the linguists at the University of Calgary: the faculty of the Linguistics Division, the graduate students of *A Higher Clause*, and the undergraduate students of *Verbatim* for listening to me rant about these patterns for nearly half a decade. Finally, I especially thank my supervisor, Darin Flynn, for his time, energy, and thoughts that helped me develop this analysis and this paper.

Territorial acknowledgements

I prepared this paper from the area where the Bow and Elbow rivers meet in Southern Alberta, a place known in Blackfoot as Mohkinstsis, a name now used to refer to the city of Calgary. This land lies within the territory of Treaty 7, which was signed by this land's stewards, representatives of the Blackfoot Confederacy, the Tsuut'ina First Nation, and the Stoney Nakoda. Additionally, this land is home to speakers of Cree and Ktunaxa, as well as members of the Métis Nation of Alberta.

I am a newcomer to this land, coming from Bulbancha, the land of many tongues at the mouth of the Mississippi, the Great River. I strive to learn more about the land, its peoples, their languages, and their history every day.

This research also took place in Iximulew (Guatemala) the traditional territory of the Kaqchikel, between Lake Atitlán, Armita (Guatemala City) and Junajpu' (Volcán de Agua). I am thankful to have been welcomed into those lands to learn and to share within the ways of knowing and living of the Kaqchikel Maya people. The data analyzed come from my fieldwork that took place during my stay there with them.

References

AnderBois, Scott. 2011. Strong positions and laryngeal features in Yukatek Maya. In Suzi Lima, Kevin Mullin & Brian Smith (eds.), *North East Linguistic Society (NELS) 39*, vol. 1, 41–54. GLSA.

Avery, Peter & William J. Idsardi. 2001. Laryngeal dimensions, completion and enhancement. In T. Alan Hall (ed.), *Studies in Distinctive Feature Theory*, 41–70. Moulton de Gruyter.

Bennett, Ryan. 2016. Mayan phonology. *Language and Linguistics Compass* 10(10). 469–514.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge & New York: Cambridge University Press.

Blevins, Juliette. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32. 117–166.

Brown, R. McKenna, Judith M. Maxwell & Walter E. Little. 2006. *¿La ütz awäch? Introduction to Kaqchikel Maya language*. Austin: University of Austin Press.

Campbell, Lyle, Terrence Kaufman & Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62(3). 530–570. DOI: 10.1353/lan.1986.0105.

Cho, Taehong. 2016. Prosodic boundary strengthening in phonetics-prosody interface. *Language and Linguistics Compass* 10(3). 120–141. DOI: 10.1111/lnc3.12178.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2. 225–252. http://www.jstor.org/stable/4419958.

Davis, Stuart & Mi-Hui Cho. 2003. The distribution of aspirated stops and /h/ in American English and Korean: An alignment approach with typological implications. *Linguistics* 41(4). 607–652. DOI: 10.1515/ling.2003.020.

Dayley, Jon P. 1985. *Tzutujil grammar*, vol. 107 University of California Publications in Linguistics. Berkeley: University of California Press.

Dresher, Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Elfner, Emily J. 2006. *The mora in Blackfoot*. University of Calgary. (MA thesis).

Fallon, Paul D. 2002. *The synchronic and diachronic phonology of ejectives*. London: Routledge.

Flynn, Darin, Merion Hodgson & Joseph W. Windsor. 2019. Aspiration in Blackfoot and Plains Cree. In *Annual Conference of the Canadian Linguistic Association*.

Guatemala, Congreso de la República de. 2003. *Ley de idiomas (law of languages) 2003*. https://www.congreso.gob.gt/assets/uploads/info_legislativo/decretos/19-03.pdf.

Guatemala, Presidencia de la República de. 1987. *Agreement recognizing the orthography for the Mayan languages of Guatemala*. https://leyes.infile.com/index.php?id=182&id_publicacion=5687.

Hall, Daniel Currie. 2007. *The role of representation of contrast in phonological theory*. University of Toronto. (Doctoral dissertation).

Halle, Morris & Kenneth N. Stevens. 1971. A note on laryngeal features. *Quarterly progress report, Research Laboratory of Electronics, MIT* 101. 198–212.

Hamann, Silke & Veronica Miatto. 2024. Three language-specific phonological interpretations of release bursts and short vowel-like formants. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 199–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264544.

Harris, John & Geoff Lindsey. 1995. The elements of phonological representation. In Jacques Durand & Francis Katamba (eds.), *Frontiers of phonology: Atoms, structures, derivations*, 34–79. Moulton de Gruyter.

Heaton, Raina & Igor Xoyón. 2016. An assessment of linguistic development in a Kaqchikel immersion school. *Language Documentation & Conservation* 10. 497–521. http://hdl.handle.net/10125/24716.

Iosad, Pavel. 2016. Prosodic structure and suprasegmental features: Short-vowel stød in Danish. *The Journal of Comparative Germanic Linguistics* 19. 221–268.

Itô, Junko, Armin Mester & Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26(4). 571–613. https://www.jstor.org/stable/4178915.

Iverson, Gregory K. & Joseph C. Salmons. 2007. Domains and directionality in the evolution of German final fortition. *Phonology* 24(1). 121–145. DOI: 10.1017/s0952675707001133.

Kehrein, Wolfgang & Chris Golston. 2004. A prosodic theory of laryngeal contrasts. *Phonology* 21. 1–33. DOI: 10.1017/S0952675704000302.

Keyser, Samuel J. & Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1). 33–63. DOI: 10.1353/lan.2006.0051.

Ladefoged, Peter. 1971. *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter. 1983. The linguistic use of different phonation types. In Diane M. Bless & James H. Abbs (eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 351–360. College Hill Press.

Launey, Michel. 1992. *Introducción a la lengua y a la literatura nahuatl*. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas,

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas: SIL. https://www.ethnologue.com/language/cak.

Lombardi, Linda. 1994. *Laryngeal features and laryngeal Neutralization*. London: Routledge.

MacKay, Carolyn. 1994. A sketch of Misantla Totonac phonology. *International Journal of American Linguistics* 60(4). 369–419. DOI: 10.1086/466242.

Montgomery-Anderson, Brad. 2008. *A reference grammar of Oklahoma Cherokee*. University of Kansas. (Doctoral dissertation).

Moseley, Christopher (ed.). 2010. *Atlas of the world's languages in danger*. 3rd edn. Paris: Unesco. http://www.unesco.org/culture/en/endangeredlanguages/atlas.

Nasukawa, Kuniya, Phillip Backley, Yoshiho Yasugi & Masatoshi Koizumi. 2018. Challenging cross-linguistic typology: Right-edge consonantal prominence in Kaqchikel. *Journal of Linguistics* 55. 1–31. DOI: 10.1017/S0022226718000488.

Nelson, Brett C. 2020. Right-edge phonological phenomena in Kaqchikel. *Calgary Working Papers in Linguistics* 31. https://prism.ucalgary.ca/handle/1880/112801/.

Patal Majzul, Filiberto, Lolmay Pedro Oscar García Matzar & Ixchel Carmelina Espantzay Serech. 2000. *Rujunamaxik ri Kaqchikel chi' : Variación dialectal en Kaqchikel*. Ciudad de Guatemala: Cholsamaj.

Ridouane, Rachid. 2006. *On the feature [spread glottis]*. DOI: 10.1.1.504.3088.

Rill, Justin. 2013. Vowel tensing in Kaqchikel Maya. In *37th Penn Linguistics Colloquium*.

Rubin, Edward J. & Aaron Kaplan. 2024. Segmental and prosodic influences on Bolognese epenthesis. In Ji Yea Kim, Veronica Miatto, Andrija Petrović & Lori Repetti (eds.), *Epenthesis and beyond: Recent approaches to insertion in phonology and its interfaces*, 105–121. Berlin: Language Science Press. DOI: 10.5281/zenodo.14264536.

Vaux, Bert. 1998. The laryngeal specification of fricatives. *Linguistic Inquiry* 29(3). 497–511. https://www.jstor.org/stable/4179030.

Warkentin, Viola & Ruth Brend. 1974. Chol phonology. *Linguistics* 12(132). 87–102.

Wax Cavallaro, Maya C. 2020. [SG] and final consonant allophony in Tz'utujil. In Ryan Bennett, Richard Bibbs, Mykel Loren Brinkerhoff, Max J. Kaplan, Stephanie Rich, Amanda Rysling, Nicholas Van Handel & Maya C. Wax Cavallaro (eds.), *Supplemental Proceedings of the 2020 Annual Meeting on Phonology*. Linguistic Society of America. http://journals.linguisticsociety.org/proceedings/index.php/amphonology/index.

Windsor, Joseph W. 2016. Contrast, phonological features, and phonetic implementation: Aspiration in Blackfoot. *Calgary Working Papers in Linguistics* 29. 61–80. https://prism.ucalgary.ca/handle/1880/51784.

Windsor, Joseph W. & J. L. Cobler. 2013. A unified analysis of three phrase-final phenomena in Blackfoot. In *Annual Conference of the Canadian Linguistics Association*.

Name index

Epenthesis and beyond

The study of epenthesis, or the insertion of a non-etymological segment, has been at the core of phonological theory from the start, and recent approaches extend beyond phonology to include phonetic considerations, as well as morphological, morphosyntactic, and lexical interactions. This volume includes 12 of the many papers presented at the workshop "Epenthesis and Beyond" held at Stony Brook University in 2021, whose goal was to provide a forum for scholars who approach epenthesis and other types of insertion from new perspectives. The articles selected for this volume represent the exciting new approaches to epenthesis that linguists are engaged in. They cover a wide range of research questions, including how different types of insertion within the same language can use different epenthetic segments, and how across languages the same phonetic material can have different phonological interpretations. Topics like feature epenthesis, insertion vs. deletion, vowel predictability, nucleus-less syllables, and epenthetic segment quality, are also explored. Some of the new tools employed by the authors include ultrasound, Information Theory, and textsetting (the study of the way poets map their text onto a metrical grid). The breadth of languages investigated is noteworthy as well: Kru languages (spoken in Western Africa), Anindilyakwa (spoken in Australia), Yuman languages (spoken in the border area between Mexico, California, Arizona), Motu (Oceanic language spoken in Papua New Guinea), Kaqchikel (Mayan language spoken in Guatemala), Arabic, Turkish, Korean, and many others.