VOWEL DELETION AS GRAMMATICALLY CONTROLLED GESTURAL OVERLAP IN USPANTEKO

RYAN BENNETT
University of California, Santa Cruz

ROBERT HENDERSON
University of Arizona

MEG HARVEY
University of Arizona

Uspanteko is an endangered Mayan language spoken in Guatemala. Unstressed vowels in Uspanteko often delete, though deletion is variable within and across speakers. Deletion appears to be phonological: it is sensitive to foot structure, morphology, and certain phonotactics, and it occurs in slow, careful speech. But deletion also has characteristics more typical of a phonetic process: it is intertwined with a pattern of gradient vowel reduction and is insensitive to most phonotactics. Electroglottography data shows that even ‘deleted’ vowels may contribute voicing to \( [\zeta(V)\zeta] \) intervals when flanked by voiceless consonants. This suggests that ‘deleted’ vowels are represented in the input to speech production, even when they are acoustically masked by articulatory overlap with adjacent segments. We conclude that vowel deletion is grammatically controlled gestural overlap, consistent with the claim that phonological representations encode information about the relative timing and coordination of articulatory gestures (e.g. Browman & Goldstein 1986, Gafos 2002). At a minimum, language-specific phonetic processes must have access to more fine-grained, abstract grammatical information than is usually assumed.*

Keywords: Mayan, articulatory phonology, gestural coordination, syncope, vowel deletion, phonetics-phonology interface

1. INTRODUCTION: THE PHONETICS-PHONOLOGY INTERFACE. Phonological theories have historically centered on the analysis of discrete units, like segments or features, which serve as the basis of phonological contrasts, alternations, and constraints. These discrete, abstract units must also be linked with a corresponding phonetic realization, in continuous time and physical space, during actual acts of speaking or listening. Given an abstract symbolic representation like \( \text{prints} \) \([p\sight\!\text{\`rints}]\), what principles relate that representation to the articulatory and acoustic variation responsible for physical phonetic forms like \([p\sight\!\text{\`rints}]\) (e.g. Hockett 1955)? Research at the PHONETICS-PHONOLOGY INTERFACE is often concerned with this notoriously difficult question.

Most work in generative phonology assumes that underlying lexical representations like \(/p\sight\!\text{\`int-z}/\) are first mapped to surface phonological outputs, like \([p\sight\!\text{\`intz}]\), which include predictable allophonic details (e.g. aspiration, voicing assimilation) along with prosodic structure (e.g. stress and syllabification). These representations are symbolic, abstract, and categorical: they do not encode information about the physical, phonetic

* We are very grateful to the Uspanteko speakers who have taught us about their language over the years, or have otherwise contributed to our research. We would especially like to thank Alejandro Pedro Vázquez Tay, Salvador Pinula Ical, Rosa Lidia Ajpoop, Juana Bernadina Ajpop Tiquiram, Tomás Alberto Méndez López, and the Comunidad Lingüística Uspanteka. K’omo chawe chaq! For helpful feedback on this project, we thank audiences at UC San Diego, UC Berkeley, UC Santa Cruz, MIT, Form and Analysis in Mayan Linguistics VI, and the Keio-ICU Linguistics Colloquium Series, especially Amanda Rysling, Jaye Padgett, John Kingston, Shigeto Kawahara, Hannah Sande, Sharon Rose, Marc Garellek, Jason Shaw, Donca Steriade, Edward Flemming, Adam Albright, Jennifer Bellik, Richard Bibbs, and Maya Wax Cavallaro. Comments from two anonymous referees and the editors at Language led to substantial improvements in the form and content of this article. This material is based upon work supported by the National Science Foundation under Grant Nos. BCS/DEL-1757473 (to Bennett) and BCS/DEL-1551666 (to Henderson). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

399

Printed with the permission of Ryan Bennett, Robert Henderson, & Meg Harvey. © 2023.
realization of speech sounds in space and time, apart from a coarse specification of linear order.

Early proposals in generative phonology assumed that these surface phonological representations are then mapped to physical phonetic outputs according to principles of phonetic implementation that are uniform across languages (Chomsky & Halle 1968:Ch.7). Subsequent research, however, has demonstrated substantial crosslinguistic variation in the phonetic realization of segments that, by phonological criteria, should be considered ‘the same’ (Keating 1984b). For example, Cho and Ladefoged (1999) report that the mean voice onset time (VOT) of contrastive, phonemic /kʰ/ is 84 ms in Hupa, but 154 ms in Navajo. This is not an isolated finding: comparable variation occurs in crosslinguistic patterns of coarticulation (Öhman 1966, Keating 1990b, Solé 1992, 1995, 2007, Manuel 1999), vowel and consonant duration (Lehiste 1970, Keating 1984b, Ladefoged & Maddieson 1996, Tang & Harris 2014, Bennett, Harvey, et al. 2022), stop release (Steriade 1994, Zsiga 2000), constriction formation (Hamann 2003), preboundary lengthening (Paschen et al. 2022), and other physical aspects of speech production. Similar observations have been made in the domain of speech perception: the acoustic cues that listeners rely on to distinguish phonemic contrasts like /s f/ vary from language to language, depending on the overall system of phonological contrasts (Wagner et al. 2006, Cutler 2012). The nonuniformity of these phonetic patterns indicates that the mapping between surface phonological outputs and speech articulation or perception must be carried out by mechanisms that are at least partly language-specific—and therefore learned—rather than fully universal.

Language-specific phonetic patterns implicate a degree of phonetic planning, controlled and implemented by speakers in much the same way that language-specific phonological patterns are (see especially Keating 1984b, Kingston & Diehl 1994). To illustrate with another example: coarticulation for nasality in vowel-nasal sequences is greater in English than it is in Spanish (Solé 1992, 1995, 2007). This entails that coarticulation and other fine details of phonetic patterning are controlled and planned by speakers—they are not merely mechanical by-products of producing segments in a particular sequence (Öhman 1966, Whalen 1990). Consequently, a full theory of learned sound patterns must account for the fact that gradient phonetic behavior, like coarticulation, can be learned and language-specific. In other words, a theory of LANGUAGE-SPECIFIC PHONETIC PLANNING is required (Kingston & Diehl 1994).

Phonetic planning has often been modeled by means of PHONETIC REPRESENTATIONS that are regulated by a language-specific PHONETIC GRAMMAR of some kind (e.g. Zsiga 2000; see also Zsiga 2021 and citations below). Phonetic representations are richer than surface phonological representations, as they include information about the realization of speech sounds in time and space, at least in an abstract and schematic form. Depending on the particular theory involved, phonetic representations may encode the relative timing of segments and articulatory gestures (e.g. Gafos 1999, 2002, Zsiga 2000), the relative magnitudes of articulatory gestures (e.g. Kingston 1984, Jun 1996), and the acoustic consequences of particular articulations (e.g. Flemming 1995, 2001, Steriade 2001, 2009, Boersma 2011, Byun et al. 2016). The phonetic grammar implements controlled,  

1 The production of phonetic and phonological patterns during speech is largely unconscious, just like many other aspects of linguistic behavior (e.g. Chomsky 1965). We use terms like ‘intentional control’, ‘cognitive control’, and ‘grammatical control’ to refer to phonetic and phonological processes that are in some sense PLANNED by speakers, albeit unconsciously and implicitly, and that must therefore be learned (Kingston & Diehl 1994, Solé 2007).
language-specific phonetic patterns in terms of these representations: the VOT difference for Hupa /kʰ/ vs. Navajo /kʰ/, for example, can be modeled as a difference in the relative timing and/or magnitude of the laryngeal gesture associated with aspiration (Goldstein & Browman 1986). Eventually, these phonetic representations are realized as actual events of speaking and/or listening, possibly according to general principles of motor control and audition (e.g. Saltzman & Munhall 1989, Delgutte 1997).

The preceding discussion assumes, at least implicitly, that phonetic representations are distinct from phonological representations (e.g. Keating 1988, 1990a). This is the position taken by modular theories of the phonetics-phonology interface. In modular theories, phonological representations are abstract and categorical, as described above. Phonological forms are then mapped to distinct phonetic representations, which are defined in continuous, physical terms (e.g. Keating 1984a, b, 1990a, Zsiga 1997, 2000, Boersma 2011). In such theories, an abstract, static form like [pʰunt] must be converted into a time-varying motor plan for speech production and must be recoverable during speech perception through some decoding of a time-varying acoustic signal.

Nonmodular theories of the phonetics-phonology interface instead assume that there is no meaningful distinction between phonetic and phonological representations. In such theories, the units of phonological representation are the same as the units of phonetic representation, be they articulatory plans (e.g. Browman & Goldstein 1986, 1992a, Pouplier 2011) or acoustic specifications (e.g. Flemming 2001). Phonological contrasts, constraints, and patterns of alternation are thus stated over representations that include at least some information about the physical realization of speech sounds in time and space. Categorical phonological behavior, such as the contextual neutralization of contrasts, is treated as the extreme endpoint of gradient behavior along a phonetic scale (e.g. Steriade 2000, 2001, Flemming 2001, Gafos & Beňuš 2006, Pierrehumbert 2016; see Zsiga 2021 for an overview). Such frameworks are sometimes known as integrated theories because they treat phonological and phonetic patterns as two sides of the same coin, instead of assigning them to different subsystems of the grammar.

This article provides a case study of grammatically controlled phonetic knowledge in Uspanteko, a Guatemalan Mayan language. Our focus is a pervasive pattern of vowel deletion in prosodically weak positions, as in /q-ʃim/ → [ˈqíʃm] ‘our corn’. We provide evidence that vowel ‘deletion’ in Uspanteko does not involve the literal elimination of a vowel from phonological representations. Instead, it is an extreme case of gradient, phonetic vowel reduction: a shortened and diminished vowel articulation overlaps with adjacent consonants to such an extent that the vowel is rendered inaudible, despite being weakly articulated (e.g. Browman & Goldstein 1992a). At the same time, these patterns of reduction and overlap show clear signs of grammatical conditioning. Vowel deletion is sharply restricted by phonological and morphological factors and does not depend on speech rate or style, occurring frequently in slow, careful speech. Vowel reduction and deletion in Uspanteko thus have the profile of phonetic processes that are regulated by an abstract, categorical grammar.

a confirmation of the predictions of such theories. If instead phonetics and phonology are distinct modules, with vowel deletion belonging to the phonetic component of the grammar, then phonetic planning must make direct reference to fairly abstract phonological and morphological structures. In either case, the boundary between phonetics and phonology would appear to be blurrier than is sometimes assumed (Myers 2000, Scobbie 2007).

In this article, we construct an analysis of vowel reduction and ‘deletion’ in Uspanteko within the framework of articulatory phonology. We are neutral as to whether this analysis should be construed as belonging to a process of phonetic implementation, as in modular theories (e.g. Zsiga 1997), or to the phonology proper, as in integrated theories (e.g. Gafos 1999, 2002). In developing our analysis, we assess several diagnostics for distinguishing phonetic and phonological patterns. In doing so, we find that some standard diagnostics for phonologically controlled behavior are in fact equivocal as to the grammatical status of vowel deletion in Uspanteko.

2. USPANTEKO. Uspanteko is an endangered Mayan language, spoken by up to 6,000 people in the central highland region of Guatemala (Richards 2003, Bennett, Harvey, et al. 2022, Us Maldonado n.d.(b)). It belongs to the K’ichean branch of the Mayan family, along with better-studied languages like K’iche’ and Kaqchikel. The grammatical structure of Uspanteko has a number of characteristics that clearly set it apart from other languages within the K’ichean branch (including lexical tone, discussed below). Unfortunately, many children in traditionally Uspanteko-speaking areas are now growing up with K’iche’ and/or Spanish as their primary languages. For more information on Uspanteko language and culture, see Can Pixabaj 2007, Us Maldonado 2010, n.d.(b), Bennett, Harvey, et al. 2022, Bennett, Henderson, & Harvey 2022, Henderson et al. 2022, and references there.

We have carried out regular fieldwork with Uspanteko speakers in Guatemala since 2010. The generalizations that we present here are based on extensive data collection over the last decade, and they reflect patterns that occur both in structured elicitation and in more spontaneous speech. Our characterization of these patterns converges with prior descriptions of Uspanteko, as well as other related Mayan languages (Campbell 1977, Barrett 1999, Can Pixabaj 2007).

3. THE PROSODY OF USPANTEKO. The word-level prosody of Uspanteko is described and analyzed in Bennett & Henderson 2013, Bennett, Harvey, et al. 2022, and Bennett, Henderson, & Harvey 2022, which build on earlier descriptions in Grimes 1972, Kaufman 1976, Campbell 1977, and Can Pixabaj 2007. Uspanteko is one of the few Mayan languages to have innovated a system of lexical tone, and it is the only such language spoken in Guatemala (Bennett 2016, England & Baird 2017, DiCanio & Bennett 2021, Bennett, Henderson, & Harvey 2022). Bennett and Henderson (2013) analyze the tone system as a privative [H] ~ ∅ contrast on stressed syllables, and we adopt this analysis here (see also Bennett, Henderson, & Harvey 2022).

Tone interacts with stress placement in Uspanteko. In toneless words, stress falls on the final syllable, as in 1. There is no secondary stress.3

2 Cunenteco, a Mayan language spoken near Uspantán, appears to have lexical tone as well (Perry Wong, p.c. 2018; see also Can Pixabaj 2017). However, there are no published descriptions of tone in this language (which is normally characterized as a dialect of K’iche’).

3 Examples in this article are primarily from our own fieldwork, carried out with speakers of Uspanteko during annual trips to Guatemala between 2010 and 2019, and remotely during 2020–2021. This fieldwork
(1) Default word-final stress
a. tinach’ab’eej [ti.na.ʃa.ˈɓeχ] ‘talk to me!’
b. ajq’ojom [ʔaχ.qˈo.ˈχom] ‘marimba player’
c. ak’ajool [ʔa.ʃa.ˈχoɡ] ‘your son (of a man)’

Vowel length is contrastive in Uspanteko (e.g. k’am [kˀam] ‘bring (it)!’ ~ k’aaam [k’am] ‘twine’), but long vowels (and vowel length contrasts) are restricted to word-final stressed syllables.

In words with long vowels, [H] tone may appear on the stressed, final syllable containing that long vowel, as in (2).

(2) Tonal long vowels: […]V
a. xáab’ [ʃá.ˈɓ] ‘vomit’
b. acháaj [ʔa.ʃá.ˈχ] ‘your ash’
c. póotz’ [ˈpó.ts] ‘blind’
d. ink’áaj [ʔin.ˈká.χ] ‘my flour’

Stress and tone are phonetically separable in Uspanteko: stress is primarily cued by duration, intensity, and vowel quality, while tone is primarily cued by raised f0 (Bennett & Henderson 2013, Bennett, Harvey, et al. 2022, Bennett, Henderson, & Harvey 2022).

In words with only short vowels, [H] tone occurs on the penultimate syllable as in (3), rather than the final syllable. Additionally, stress retracts to the penult to coincide with tone.

(3) Penultimate accent with tonal short vowels: […]Vς
a. inkínaq’ [ʔin.ˈkí.naq舒心] ‘my bean’
b. ajxójil [ʔay.ʃó.χil] ‘dancer’
c. wersálik [wer.ˈsá.lik] ‘asleep’
d. muqunélib’ [mu.qu.ˈné.lিব舒心] ‘gravediggers’

These patterns reflect two generalizations: [H] tone, when present, always falls on the penultimate mora of the word (either […]Vςς (3) or […]Vςμ (2)); and stress and tone always occur on the same syllable.

Bennett and Henderson (2013) analyze tone and stress placement in Uspanteko in terms of foot structure. Default word-final stress, as in 1–2, involves an iambic foot at the right edge of the word, […](ςς). Penultimate accent, as in 3, amounts to a change in foot type: the foot remains anchored at the right edge, but becomes trochaic rather than iambic, […](ςς). This change in foot type ensures that stress and tone occur on the same syllable, while also respecting the invariant placement of [H] tone on the penultimate mora.

Confirmatory evidence for this analysis comes from the interaction of stress assignment with vowel deletion, which we now turn to.

3.1. Vowel deletion. The deletion of unstressed short vowels is a salient feature of spoken Uspanteko.4 Deletion is variable but quite commonplace, giving rise to frequent alternations in the form of individual morphemes (4). Long vowels are always stressed (§3), so only unstressed short vowels participate in alternations like 4.
(4) Vowel deletion with default final stress

a. masaat [ma.'sa:t] ~ [m.'sa:t] ‘deer’
b. chukuy [tu.'kuy] ~ [t.'kuy] ‘pine cone’
c. chik’oor [ti.'k'or] ~ [t.'k'or] ‘hiccup’
d. richijiil [ri.ti.'chiil] ~ [ri.ti.'chiil] ‘her husband’

For reasons outlined in §§4.1–7 below we assume that vowel deletion does not change
syllabification, and hence represent the output of deletion in 4 with apparently ‘degener-
ate’ syllables like [.C.] (see also Bennett & Henderson 2013, Kawahara & Shaw 2018).

Deletion is conditioned by the position of the vowel relative to stress. In words with
final stress, deletion systematically occurs in the pretonic, penultimate syllable […σ]
Figure 1 shows variable deletion of pretonic /o/ in the word tijq’ojomaan [tiχ.qʼo.χ(o).′maːn]
‘(s)he plays the marimba’, as produced by two different speakers reading a
word list.

In words with final stress, deletion is essentially limited to pretonic position […σ].
Deletion is far less common in other positions, even in casual speech. For example,
in muquneel [mu.qu.ˈneːl] ‘gravedigger’, unstressed [u] freely deletes in pretonic posi-
tion, but not in the initial syllable. The lack of deletion in the initial syllable cannot be
attributed to segmental phonotactics, as deletion regularly produces word-initial conso-
nant sequences (e.g. masaat [ma.'sa:t] ~ [m.'sa:t] ‘deer’; see §4.7). Instead, vowel dele-
tion is selective in applying only to unstressed vowels in specific positions. We take
this asymmetry to be a fundamental property of vowel deletion in Uspanteko, which
must be accounted for under any analysis.

The locus of deletion is dependent on word-level accent. As we have seen, deletion
selectively targets the pretonic syllable in words with final accent […σ]. But in words
with penultimate accent, as in 3 above, deletion targets the posttonic syllable instead,
 […′σ] (Figure 2).

Postonic deletion is also selective: in muqunélib’ [mu.qu.'né.liɓ] ‘gravediggers’, for
instance, the postonic short vowel may freely delete, but the other two unstressed short
vowels may not.

Bennett and Henderson (2013) argue that the correlation between accent placement
and the locus of vowel deletion supports their foot-based analysis of Uspanteko pros-
yody. In words with final stress, deletion targets the pretonic syllable. Under a metri-
cal approach to stress, this is equivalent to deleting the vowel in the weak branch of
the foot, […′σ′]. This analysis extends straightforwardly to words with penultimate

\(^5\) Figure 1 is unusual in showing deletion of a mid vowel, a point we discuss in detail in §4.8.
Vowel deletion as grammatically controlled gestural overlap

accent: assuming that stress retraction involves trochaic footing [...]σ̂σ̂], posttonic vowel deletion in words like [ʔin.ˈkín(a)q'] (Fig. 2) is again simply deletion in the weak branch of the foot. Foot structure thus offers a unified account of the locus of vowel deletion in Uspanteko.

Deletion is variable within the speech of individuals. Indeed, speakers of Uspanteko will often produce the same item with and without deletion, sometimes with just a few seconds between each rendition. Figure 3 illustrates this phenomenon with the tonal word inpix [ʔímp(i)ʃ] ‘my tomato’.

Vowel deletion in Uspanteko thus involves token-wise variability: across multiple utterances, the exact same word may or may not show deletion (Fig. 3).6

4. Categorical deletion vs. articulatory overlap. The extensive variability observed for vowel deletion in Uspanteko raises an important question (Bennett & Henderson 2013:625–28): could deletion be a phonetic phenomenon, resulting from patterns of articulatory coordination during speech production, rather than belonging to the abstract, categorical phonology as such?

Many phonological processes have been described as applying optionally or variably. However, at least some of these processes are probably best understood as being phonetic in character: they are phonetically gradient, phonetically incomplete, and/or conditioned by factors that are normally irrelevant for phonological patterning (e.g. speech rate; §4.1). For example, word-final /t/, d/-deletion in American English has been characterized as an optional but categorical process, as in West Side [wes(t) # said] (see

6 Token-wise variability should be distinguished from lexical variability, which describes any process that consistently applies to some words or morphemes but not others (Zuraw 2010, 2016, Gouskova 2012, Zymet 2018, Hout 2020, among many others). Vowel deletion in Uspanteko does not show lexical sensitivity of this sort. Lexical variability could also be called type-wise variability, since it involves nonuniform behavior across different items (rather than across tokens of the same item).
Coetzee & Pater 2011 for references). If correct, this would be a variable phonological pattern (§3.1). But instrumental phonetic evidence suggests that the deletion of /t, d/ is often incomplete: the tongue-tip gestures for /t, d/ are still present, but are heavily reduced and/or acoustically masked by gestural overlap with neighboring segments (Browman & Goldstein 1986, 1989, 1990, 1992a, Purse 2019). At a minimum, this suggests that /t, d/-deletion cannot be taken as a clear case of an optional-but-categorical phonological process. Similar observations have been made for l-darkening in several varieties of English (e.g. Lee-Kim et al. 2013, Turton 2017, Mackenzie et al. 2018), schwa deletion in American English and French (Kaisse 1985, Browman & Goldstein 1992a, Barnes & Kavitskaya 2002, Davidson 2006b, Bürki et al. 2011), and nasal place assimilation in American English (Ellis & Hardcastle 2002) (see Kawahara 2011:§2.2 for other examples).

The variable processes cited above all involve gradience, which we define as continuous differences in the magnitude of some measure along a physical or temporal scale. By way of illustration, Lee-Kim et al. (2013) report that velarization, or ‘darkening’, of American English /l/ occurs along a continuum: word-final /l/ in tall is darker than preboundary /l/ in tall-est, which is in turn darker than postboundary /l/ in flaw-less (see also Sproat & Fujimura 1993, Turton 2017). This is a gradient pattern, because the degree of darkening varies continuously depending on the context /l/ occurs in.7

As Cohn (2006) and Zsiga (2021) note, physical gradience is often taken to be a property of phonetic rather than phonological processes. Hence, variable processes that also show gradience may in fact owe to phonetic processes of articulatory coordination and reduction, rather than abstract phonological rules or constraints (Browman & Goldstein 1990, 1992b, Myers 2000, Davidson 2003, 2006a,b; on /l/-darkening specifically, see Turton 2017).

To be sure, categorical phonological processes can also apply variably: for example, the devoicing of voiced geminates in Japanese borrowings like [beddo] ~ [betto] ‘bed’ is optional, but phonetically complete (Kawahara 2006; see also Cohn 1993, Zsiga 1997, 2021, Kochetov & Pouplier 2008, Zuraw 2010, Coetzee & Pater 2011:§3.1, Strycharczuk 2012, Strycharczuk & Simon 2013, Shaw & Kawahara 2018, Plug et al. 2019). Since both phonological and phonetic patterns may be variable, variability itself is not a good diagnostic for distinguishing phonological vs. phonetic processes. At the same time, the phonetic evidence outlined above suggests that variable sound patterns should be closely scrutinized, with an eye to the possibility that at least some of those patterns might in fact be phonetic in nature. This is especially true for variable processes—like vowel deletion in Uspanteko—that involve segmental reduction or assimilation, as such phenomena are often amenable to explanations based on the reduction and coordination of articulatory gestures (Browman & Goldstein 1990).


---

7 Gradience in this sense, which involves continuous differences in physical space or time, should be distinguished from variability, understood here as categorical variation between tokens or lexical items (§3.1). ‘Gradience’ is also used to describe nonphysical phenomena (e.g. intermediate or noncategorical well-formedness judgments; Daland et al. 2011), but we focus here on the physical and temporal senses of this term. See Myers 1995, Zsiga 1997, 2021, Cohn 2006, 2007, Kochetov & Pouplier 2008, Chitoran & Cohn 2009, and Zymet 2018 for discussion.
Vowel deletion as grammatically controlled gestural overlap

Goldstein (2011), Ridouane and Fougeron (2011), and Bellik (2018), among others, have all argued that certain cases of vowel epenthesis are better understood as arising from patterns of gestural overlap. Bellik (2018), for example, presents a range of evidence converging on the conclusion that ‘epenthetic’ vowels in Turkish onset clusters are not phonologically inserted, but are instead the result of low gestural overlap (i.e. an open transition) between successive consonants. Similarly, Barnes and Kavitskaya (2002) and Davidson (2006b) argue that schwa ‘deletion’ in French and English corresponds to high gestural overlap between consonants in a [C\textcircled{a}C] sequence during speech production, which phonetically masks underlying [æ], and should therefore not be analyzed as the result of categorical deletion of a vowel in the symbolic phonology (see also Browman & Goldstein 1992a, Bürki et al. 2011). Vowel devoicing, which can be understood as a species of vowel reduction closely related to vowel deletion, has also been analyzed as a consequence of high gestural overlap (Kondo 1994, 2008, Jannedy 1995, Beckman 1996, Jun et al. 1997, Delforge 2009, Dabkowski 2018).

In §4.1 we consider several diagnostics that have been proposed to distinguish phonological from nonphonological patterns. We find that some of these diagnostics are equivocal as to the grammatical status of vowel deletion in Uspanteko. Phonotactic restrictions on deletion suggest that it is phonologically conditioned, but the same patterns can be plausibly explained in phonetic terms as well. A phonetic perspective on vowel deletion also provides insight into apparent exceptions to phonotactic conditioning that are surprising from a purely phonological point of view. Perhaps most importantly, we show that deletion is the extreme endpoint of a gradient process of foot-internal vowel reduction. If physical gradience is indicative of phonetic rather than phonological patterns, this strongly implies that deletion is a phonetic process.

Ultimately, three diagnostics—involving speech rate, morphology, and prosodic structure—provide the best evidence that deletion is under grammatical control in Uspanteko, despite showing the phonetic characteristics alluded to above. Section 7 attempts to cut the Gordian knot by providing a formal analysis of vowel deletion in terms of grammatically controlled gestural overlap.


Other diagnostics—such as phonological sensitivity to lexical category distinctions like ‘noun’ vs. ‘verb’ (Myers 2000, Smith 2011)—are not relevant for the patterns we discuss, so we do not include them in Table 1.

All of the diagnostics we use to identify vowel deletion as a phonological process could be disputed, and we engage with many potential counterarguments in the text below. We also acknowledge that these diagnostics are to some extent theory-internal. For example, Flemming (2001) argues that gradient phonetic coarticulation can trigger (i.e. ‘feed’) categorical neutralization processes; if correct, this would at least complicate the application of our fifth diagnostic in Table 1 (see also Vennemann
These diagnostics are often presented in the literature as being tests for phonological vs. nonphonological/phonetic behavior. However, we emphasize that some of these diagnostics assess only whether a certain pattern is under some form of speaker control (e.g. sensitivity to speech rate; Gafos 2002, Solé 2007, Eischens 2022). Such diagnostics are thus ambivalent when it comes to distinguishing language-specific phonology from language-specific phonetics (Keating 1984b, 1990a, Kingston & Diehl 1994, Manuel 1999). In §9 we revisit the extent to which each of these diagnostics may or may not be a reliable indicator of phonological vs. phonetic patterning and, more generally, of intentional grammatical control.

4.2. Deletion is the endpoint of gradient reduction. Vowel deletion is a variable process in Uspanteko: foot-internal, unstressed short vowels are eligible for deletion, but deletion does not consistently apply. However, even when short vowels in this context escape deletion, they are typically phonetically reduced. These reduced vowels are often quite short and are audibly centralized relative to other unstressed vowels. This is especially true of vowels in posttonic position (a point we return to in §4.11). Figure 4 shows that elidable vowels are produced on a continuum of reduction, ranging from relatively long (~60 ms) to relatively short (~20 ms) tokens and culminating in the full loss of any audible vowel on the surface. (For reference, stressed short vowels are about 100 ms long in careful speech in Uspanteko; Bennett, Harvey, et al. 2022.)

Deletion therefore coexists alongside a gradient, continuous pattern of vowel reduction. This pattern of reduction occurs in exactly the same positions, and under essentially the same conditions, as deletion. A natural inference would be that deletion and reduction constitute a single process. Specifically, we hypothesize that both vowel reduction and vowel deletion stem from articulatory overlap between a vowel and its neighboring consonants.

When a vowel is coproduced with a consonant, such that their respective articulations overlap in time, the acoustic output during that period of articulatory overlap will typically sound like a consonant. This is because consonants are more constricted than vowels, and the output of the vocal tract is strongly determined by the point of most extreme constriction (e.g. Mattingly 1981, Stevens 1989, 2000, Goldstein et al. 2006, Pouplier 2011, Johnson 2012:Ch.7; see §8.1 for more discussion). Intuitively, if a stop or fricative constriction (for example) is produced anywhere in the vocal tract, the acoustic

<table>
<thead>
<tr>
<th>1. Show physical gradience</th>
<th>No</th>
<th>Yes</th>
<th>Yes: overlaps with V reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Dependent on speech rate</td>
<td>No</td>
<td>Yes</td>
<td>No: applies in slow speech</td>
</tr>
<tr>
<td>3. Sensitive to morphological structure</td>
<td>Yes</td>
<td>Possibly, but only gradually</td>
<td>Yes, categorically</td>
</tr>
<tr>
<td>4. Sensitive to phonotactic restrictions</td>
<td>Yes</td>
<td>No</td>
<td>Unclear; probably not</td>
</tr>
<tr>
<td>5. Feed/bleed phonological processes</td>
<td>Yes</td>
<td>No</td>
<td>Unclear; probably not</td>
</tr>
<tr>
<td>6. Sensitive to abstract prosodic structure (e.g. metrical feet)</td>
<td>Yes</td>
<td>Yes, but only gradually</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1. Prototypical characteristics of phonological (abstract, symbolic, and categorical) vs. phonetic (continuous, in real time and physical space) patterns.


These diagnostics are often presented in the literature as being tests for phonological vs. nonphonological/phonetic behavior. However, we emphasize that some of these diagnostics assess only whether a certain pattern is under some form of speaker control (e.g. sensitivity to speech rate; Gafos 2002, Solé 2007, Eischens 2022). Such diagnostics are thus ambivalent when it comes to distinguishing language-specific phonology from language-specific phonetics (Keating 1984b, 1990a, Kingston & Diehl 1994, Manuel 1999). In §9 we revisit the extent to which each of these diagnostics may or may not be a reliable indicator of phonological vs. phonetic patterning and, more generally, of intentional grammatical control.

4.2. Deletion is the endpoint of gradient reduction. Vowel deletion is a variable process in Uspanteko: foot-internal, unstressed short vowels are eligible for deletion, but deletion does not consistently apply. However, even when short vowels in this context escape deletion, they are typically phonetically reduced. These reduced vowels are often quite short and are audibly centralized relative to other unstressed vowels. This is especially true of vowels in posttonic position (a point we return to in §4.11). Figure 4 shows that elidable vowels are produced on a continuum of reduction, ranging from relatively long (~60 ms) to relatively short (~20 ms) tokens and culminating in the full loss of any audible vowel on the surface. (For reference, stressed short vowels are about 100 ms long in careful speech in Uspanteko; Bennett, Harvey, et al. 2022.)

Deletion therefore coexists alongside a gradient, continuous pattern of vowel reduction. This pattern of reduction occurs in exactly the same positions, and under essentially the same conditions, as deletion. A natural inference would be that deletion and reduction constitute a single process. Specifically, we hypothesize that both vowel reduction and vowel deletion stem from articulatory overlap between a vowel and its neighboring consonants.

When a vowel is coproduced with a consonant, such that their respective articulations overlap in time, the acoustic output during that period of articulatory overlap will typically sound like a consonant. This is because consonants are more constricted than vowels, and the output of the vocal tract is strongly determined by the point of most extreme constriction (e.g. Mattingly 1981, Stevens 1989, 2000, Goldstein et al. 2006, Pouplier 2011, Johnson 2012:Ch.7; see §8.1 for more discussion). Intuitively, if a stop or fricative constriction (for example) is produced anywhere in the vocal tract, the acoustic
Vowel deletion as grammatically controlled gestural overlap

Consequence will be a stop or fricative, regardless of whether the consonant in question is coproduced or overlapped with a vowel. Increasing the amount of overlap between a vowel and neighboring consonants will thus have the acoustic effect of shortening that vowel.

Increased overlap could result from a change in the relative timing of vowel and consonant articulations. For instance, if the articulation of a consonant in a VC sequence begins earlier, the vowel will be more overlapped, all else being equal. This will lead to a greater degree of (apparent) vowel shortening and reduction. In §8.2 we suggest that vowels may also be hypoarticulated—produced with smaller articulatory gestures—in foot-internal position in Uspanteko, which would contribute to reduction and overlap in this context.

Extremely high levels of gestural overlap may then result in the apparent ‘deletion’ of the vowel: the vowel is articulated, but acoustically it is entirely masked by the articulations of the flanking consonants (just as Browman and Goldstein (1992a) and Davidson (2006b) propose for the deletion of unstressed /ə/ in English). These scenarios are schematized in Figure 5, which uses open trapezoids to represent the timecourse of constriction formation in the oral tract for each segment (e.g. Gafos 2002).

![Figure 5](image)

Vowel reduction can thus be understood as the aggressive encroachment of consonant articulations on an intervening unstressed vowel. When gestural overlap is high enough, the articulation of the vowel may be completely hidden and end up inaudible. On this view,

---

8 This point is nicely illustrated by palatalized and velarized consonants, which are essentially consonants that are coproduced with an [i] or [u] articulation superimposed (e.g. Irish /biːnˠ/ ‘peak’; Bennett, Ní Chiosáin, et al. 2018). Such sounds are of course acoustically consonants rather than vowels.
there is no literal deletion at all in Uspanteko—just different degrees of gestural overlap, leading to gradient patterns of vowel reduction, up to and including reduction to silence.

This proposal makes an important prediction about vowel deletion in Uspanteko: in contexts where vowels are phonetically lengthened, deletion should be less common. Phonetically longer vowels should have less overlap with their neighboring consonants, and thus should tend to remain audible. Conversely, in contexts where vowels are phonetically shortened, deletion should be more common, because shorter vowels are more likely to be hidden by the articulations of flanking consonants. We return to this prediction in the following sections and show that it helps explain some fine-grained details of vowel deletion in Uspanteko.

4.3. **Speech rate and style.** We have suggested that vowel deletion or syncope in Uspanteko reflects patterns of articulatory coordination, rather than a phonological rule of vowel deletion. But curiously, syncope also has properties that are more typical of phonological processes than of phonetic ones. The first such property has to do with speech rate.

Foot-sensitive vowel deletion occurs frequently in slow, careful, formal speech. By way of illustration, Figure 6 shows two instances of the word lékej [ˈlék(e)x] ‘up, above’, produced seconds apart by the same speaker. This data was collected as part of an elicitation task in which the speaker was slowly and deliberately reading a word list, while wearing electrodes on his neck to monitor vocal-fold vibration (see §6). Despite the paced and somewhat artificial speech style, and the obviously abnormal circumstances of the recording session, deletion was nonetheless prevalent in this speaker’s productions.

![Figure 6. Variable deletion in lékej [ˈlék(e)x] ‘up, above’ in a formal elicitation task (speaker 6, 2019).](image)

This is completely typical of fieldwork with Uspanteko speakers: deletion is rampant, regardless of speech style and elicitation context. Indeed, deletion is often indicated in the Uspanteko orthography, suggesting that speakers consider deletion appropriate even for highly formal documents like dictionaries (e.g. one dictionary, Us Maldonado n.d.(a), writes both qaleen and qleen for [q(a).l.en] ‘things’, both k‘ayb‘al and k‘ayb ‘l for [k‘áj.ɓ(a)l] ‘market’, and so on).9

It is certainly possible that deletion is more frequent at faster speech rates in Uspanteko (indeed, the gestural analysis we develop in §8 probably predicts as much; see also Davidson 2006b). But the crucial observation here is that deletion is extremely common in slow, careful speech. Deletion is therefore not dependent on speech rate or style in Uspanteko. For that reason, deletion cannot be treated as a mechanical phonetic by-product of speaking quickly or as a fast-speech phonological rule.

9The fact that deletion is regularly included in written Uspanteko may itself be evidence that deletion is a phonological process; see, for example, Sapir 1925, 1933, Kenstowicz 1994a:1–11, Ch. 5, and Hall 2006 for discussion.
Independence from speech rate is a classic diagnostic for phonological rather than phonetic processes (Kaisse 1985, McCarthy 1986, Solé 1992, 1995, 2007, Keating 1996, Myers 2000, Gafos 2002, Davidson 2006b, Bürki et al. 2011, Strycharczuk 2012, Kilbourn-Ceron 2017, Eischens 2022, etc.). The fact that vowel deletion regularly occurs across all speech rates and styles provides a strong indication that speakers are in some sense intending to reduce and/or delete foot-internal vowels. While speech rate might influence the likelihood of deletion, the fact remains that deletion is entirely commonplace at speech rates and styles that otherwise favor hyperarticulated, careful pronunciations. This implies that deletion is under grammatical control.

4.4. Sensitivity to morphology. Vowel deletion in Uspanteko is conditioned by morphological structure in a way that suggests it is grammatically controlled. Vowel deletion freely targets vowels in roots (5–6) and suffixes (7).

(5) Vowel deletion in noun roots
   a. tz’ikin [t’s’ikin] ~ [t’s’kin] ‘bird’
   b. túkan [’túkan] ~ [’túk’n] ‘blackberry’
   c. qálaq [’qá-laq] ~ [’qá-lq] ‘our plate’
   d. inchikich [ʔin-’ʃiʃkʃ] ~ [ʔin-’ʃiʃkʃ] ‘my basket’

(6) Vowel deletion in verb roots
   a. xinsik’iij jwiich [ʃ-in-s’iʃk’iːχ ρχwiːt] ~ [ʃ-in-s’iʃk’iːχ ρχwiːt] ‘I read it’
   b. xijsik’in [ʃ-iχ-s’iχk’in] ~ [ʃ-iχ-s’iχk’in] ‘(s)he called’
   c. tinchajaj [t-in-’ʃiʃ-aχ] ~ [t-in-’ʃiʃ-aχ] ‘I care for it’

(7) Vowel deletion in suffixes
   a. ajchákib’ [ʔaχ=’ʃák-ɓ] ~ [ʔaχ=’ʃák-ɓ] ‘workers’
   b. chapálik [t’ʃa-p-ál-ik] ~ [t’ʃa-p-ál-k] ‘seized’
   c. k’áyb’al [’kˀáj-ɓl] ~ [’kˀáj-ɓl] ‘market’

In contrast, deletion never applies in prefixes in Uspanteko (8), even when those prefixes occur in contexts where deletion is normally allowed (e.g. word-initial pretonic syllables). Importantly, the blocking of deletion in prefixes cannot be reduced to phono-tactics: the clusters that would result from deletion in prefixes are otherwise licit in Uspanteko, either underlyingly, as the result of morpheme concatenation, or as the result of syncope in other morphological environments.

(8) No vowel deletion in prefixes10
   a. xinel [ʃ-in-’el] ~ *[ʃ-n-’el] ‘I left’
      cf. xnojisaj [ʃ-nɔχ-’i-ʃaχ] ‘it was filled’
   b. xajach [ʃ-’aχ-’aχ] ~ *[ʃ-’aχ-’aχ] ‘you separated it’
      cf. xjachsaj [ʃ-’aχʃ-’i-ʃaχ] ‘it was separated’
   c. xojok [ʃ-’oχ-’ok] ~ *[ʃ-’oχ-’ok] ‘we entered’
      cf. xajab’ [ʃaχ-’aχ] ~ [ʃaχ-’aχ] ‘shoe’
   d. tak’am [t-a-’k’am] ~ *[t-a-’k’am] ‘you receive it’
      cf. tk’ixib’ [t-ʃiʃ-’iʃb] ‘(s)he was ashamed’
   e. qalaaq [ qa-laq] ~ *[q-’laq] ‘our plate’
      cf. qaleen [qa’le̱n] ~ [’qe̱n] ‘things’

---

10 Morphophonological and lexical variation are widespread in the Uspanteko community, so some of the examples here have variant forms: qalaaq [ qa’laq] ‘our plate’, for instance, is also realized as tonal qálaq [’qá.laq] ~ [’qá.lq], as in 5. Similarly, xajab’ [ʃaχ-’aχ] ~ [ʃ-’aχ] ‘shoe’ may be realized as xajab’ [ʃaχ-’aχ] ~ [ʃaχ-’aχ]. This variation is orthogonal to the point that vowel deletion is inhibited in prefixes.
A ban on the deletion of prefix vowels is also found in Sipakapense, another K’ichean-branch Mayan language related to Uspanteko (Barrett 1999, 2011, Bennett 2016).

While we might appeal to a prohibition on the deletion of entire morphemes to explain cases like 8b,d, this explanation is not available for the remaining forms in 8. Nor can we appeal to a prosodic difference between prefixes and suffixes that might condition deletion, because the available evidence from stress, tone, and segmental phonology indicates that prefixes and suffixes are equally well integrated with their stems, a point we return to shortly below (Bennett & Henderson 2013, Bennett 2018, Bennett, Harizanov, & Henderson 2018).

Prefixes are mostly inflectional in Uspanteko, as in other Mayan languages (Coon 2016, Bennett 2018). Still, the fact that deletion is banned in prefixes cannot be reduced to a ban on deletion in inflectional morphemes. Tellingly, the inflectional plural suffix -ib'/iɓ/ regularly undergoes vowel deletion (7a above), while inflectional prefixes do not. We conclude that vowel deletion is indeed inhibited in prefixes, but not in other morphological environments.

Morphological conditioning, particularly at this level of granularity, is the hallmark of grammatically controlled processes and has been used to distinguish phonological from phonetic phenomena (Pierrehumbert 1990, 2002, Myers 2000, Bermúdez-Otero 2015, Turton 2017, and references there). The inhibition of vowel deletion in prefixes thus counts as additional evidence that deletion is under grammatical control.

**Morphological conditioning cannot be reduced to phonetic factors.** In this section we consider whether the prohibition on syncopating vowels in prefixes is reducible to phonetic factors. Our conclusion is a negative one: given our current understanding of the relationship between gestural organization and morphological structure, there is no independent reason to expect syncope to be inhibited in prefixes in Uspanteko. This is an important claim, inasmuch as it provides evidence that syncope is directly conditioned by the morphological identity of the segments involved. It follows that patterns of gestural overlap must have access to abstract grammatical information, such as the prefixal affiliation of certain vowels.

There is some evidence that speech articulation is sensitive to morphological structure. For example, Seyfarth et al. (2018) report durational differences between morphologically complex English words like *free-s* and otherwise homophonous roots like *freeze*. Oh and Redford (2012) argue that sequences of identical consonants are phonetically different across morpheme boundaries [Cₐ-C₆] vs. across word boundaries [Cₐ ≠ C₆] in English. Relatedly, Cho (2001) reports that gestural coordination in phoneme sequences may vary as a function of morphological context. Cho found that gestures were less overlapped within words than across word boundaries in Korean (see also Browman & Goldstein 1988). Similarly, intergestural timing was found to be less variable inside words than between words. Morphological effects on gestural overlap are particularly relevant for our analysis of Uspanteko, which treats syncope as extreme gestural overlap (see also Browman & Goldstein 1990, Baranowski & Turton 2020; for further examples and critical review, see Lee-Kim et al. 2013, Turton 2017, Tang & Bennett 2018, Strycharczuk 2019, Bell et al. 2021, Mousikou et al. 2021).

The literature on morphologically conditioned phonetic effects has entertained various explanations for phonetic differences related to morphological structure. To our knowledge, none of these explanations predicts a core property of syncope in Uspanteko: that it applies in stems and suffixes, but not in prefixes (5–9).
Vowel deletion as grammatically controlled gestural overlap

(9) No deletion in prefixes

a. qasuut’  [qa-ˈsu:tˀ] ~ *[q-ˈsu:tˀ]  ‘our napkin(s)’
b. xayol  [ʃ-a-ˈjol] ~ *[ʃ-ˈjol]  ‘you spoke’

As noted above, this asymmetry cannot be reduced to a prosodic difference between prefixes and suffixes. Prefixes in Uspanteko are phonologically well integrated with their hosts. They can bear stress and tone (e.g. ápix [ˈʔá-piʃ] ‘your tomato’), and they interact with their stems in phonological processes like word-initial glottal stop insertion (e.g. aqan [ʔa.ˈqan] ‘leg(s)’ vs. awáqan [ʔa.ˈw-á.qan] ‘your leg(s)’ (§4.5); see Bennett, Harvey, et al. 2022). Suffixes show similar behavior—for example, they are stressable in examples like iporii [ʔ-por-ˈiːl] ‘its roughness’. The fact that prefixes can bear stress and tone is especially important, as it demonstrates that they can be parsed within the metrical foot—the domain of deletion.

Other ways of classifying affixes—inflectional vs. derivational, productive vs. unproductive, frequent vs. infrequent, and so on—also fail to pick out prefixes as a class in Uspanteko (and in other Mayan languages; Coon 2016, Polian 2017, Bennett 2018). Such categories simply make the wrong cut.

The resistance of prefixes to syncope may reflect a historical origin as free-standing words or clitics, which eventually eroded into prefixes (Robertson 1992, Mora-Marín 2021). Synchronically, the ban on syncopating vowels in prefixes appears to be arbitrary. Arbitrariness is a sign of grammaticization, consistent with the view that vowel deletion in Uspanteko is either a phonological process or a language-specific phonetic process conditioned categorically by morphological structure.

We now turn to apparent phonotactic restrictions on syncope. These restrictions seem to suggest that syncope is a grammatically controlled process. However, each phonotactic condition we consider can also be explained in phonetic terms, often with some explanatory gain. Our conclusion is that phonotactic restrictions are at best equivocal as to the grammatical status of syncope in Uspanteko, and may in fact support a more phonetically oriented analysis.

4.5. Clusters containing /ʔ/. Vowel deletion is positionally restricted in Uspanteko. Vowel deletion never targets underlyingly vowel-initial words, which surface with an epenthetic [ʔ], /#V…/ → [#ʔV…], as seen in 10 (Can Pixabaj 2007, Bennett 2016, 2018, Bennett, Harvey, et al. 2022). This restriction is again shared with the related language Sipakapense (Barrett 1999, 2011).12

(10) No deletion in initial /#V/ → [#ʔV]

b. ab’aj  [ʔa.ˈbaʃ] ~ *[(ʔ).ˈbaʃ]  ‘stone’

Parallel to the ban on deletion in initial /#V…/ → [#ʔV…], deletion never targets word-final /Vʔ#/ sequences, as in 11. Here, /ʔ/ is phonemic and underlying rather than inserted.

11 Possible exceptions are the noun-forming agentive prefix aj- [ʔaχ=] and feminine prefix ix- [ʔiʃ=]; see Bennett 2016, 2018, and Bennett, Harizanov, & Henderson 2018.

12 Evidence that these words are underlyingly vowel-initial comes from [ʔ] ~ ∅ alternations in morphologically related pairs like okox [ʔo.ˈkoʃ] ‘mushroom’ vs. wokox [w-o.ˈkoʃ] ‘my mushroom’. See Bennett 2018 and Bennett, Harvey, et al. 2022 for extensive discussion.
(11) No deletion in final /Vʔ#/  
   a. ínka’ [ʔín.kaʔ] ~ *[ʔín.kʔ] ‘my grinding stone’  
   b. qátz’i’ [qá.ʔísʔʔ] ~ *[qá.ʔísʔʔ] ‘our dog’  
   c. ínq’u’ [ʔín.ʔuʔ] ~ *[ʔín.ʔuʔ] ‘my jacket’  

These restrictions may be reducible to a ban on [#ʔC] and [Cʔ#] clusters, which are not otherwise attested at word edges in Uspanteko. If so, the fact that syncope is inhibited in these environments might count as an instance of phonotactic sensitivity. And if phonotactic sensitivity is diagnostic of phonological processes, these restrictions would indicate that vowel deletion belongs to the symbolic phonology.

A phonetic account of [#ʔC] and [Cʔ#] clusters. If syncope involves high degrees of gestural overlap, rather than true deletion, then the ban on syncopating vowels in word-initial [#ʔVC] or word-final [CVʔ#] sequences cannot be treated as a ban on [ʔC] and [Cʔ] clusters at word edges. If syncope amounts to extreme gestural overlap, then derived /ʔVC/ → [ʔC] and /CVʔ/ → [Cʔ] ‘clusters’ are still phonologically [CVC] sequences. Hence, they do not violate constraints against [ʔC] or [Cʔ] clusters as such. Why, then, would syncope be blocked in these contexts?

A solution to this puzzle can be found if we consider the potential acoustic consequences of gestural overlap between a vowel and neighboring glottal stop [ʔ]. In Uspanteko, as in many Mayan languages, a glottal stop is realized sometimes as full glottal closure and sometimes as creaky voice on adjacent vowels and sonorants (Frazier 2009, Baird 2011, Baird & Pascual 2011, Bennett 2016, Bennett, Harvey, et al. 2022). These two phonetic realizations of [ʔ] are illustrated in Figure 7 with word-medial [ʔ], where variability between full and partial glottal closure is clearest.

We assume that creaky realizations of [ʔ] reflect coarticulation with neighboring voiced segments: when [ʔ] is overlapped with a voiced vowel or consonant, it is realized as creaky voice rather than a full stop (see also Borroff 2007, Davidson 2021, and references there). This is consistent with the fact that transitions into and out of [ʔ] are often creaky, even when [ʔ] is realized with full glottal closure (Fig. 7). Creaky realizations of [ʔ] may also stem from incomplete glottal closure in prosodically weak contexts, like unstressed syllables, where articulatory movements tend to be reduced.

These observations are key for understanding the apparent lack of deletion in [#ʔVC] and [CVʔ#] sequences. Since glottal stop lacks an oral articulation, the acoustic consequences of overlap between [ʔ] and a vowel depend entirely on how the laryngeal gesture for [ʔ] is realized. If [ʔ] is realized as a full stop, the vowel will be inaudible, given the lack of airflow through the vowel tract. But if [ʔ] is instead realized as creakiness, then the vowel will remain audible despite extensive overlap with [ʔ] (Figure 8).
Vowel deletion as grammatically controlled gestural overlap

This is what we propose for Uspanteko. In syncope contexts, [ʔ] overlaps with vowels to the same extent that other consonants do. However, [ʔ] lacks an oral articulation and is typically realized as creaky voice during periods of overlap with voiced segments in Uspanteko. As a consequence, overlap with [ʔ] does not lead to the auditory impression of vowel deletion. In other words, [#ʔVC] and [CVʔ#] sequences have exactly the same phonology and gestural organization as other syncopated [CVC] sequences, but the acoustic effects of gestural overlap are quite different. Rather than being obscured, the vowel is simply realized as creaky, as seen in Figure 9 for [CVʔ#]. Additionally, [ʔ] may be gesturally reduced in prosodically weak positions, which may further favor creaky-voiced realizations in foot-internal unstressed syllables (Garellek 2013, Davidson 2021).

We conclude that the lack of vowel deletion in word-initial [#ʔVC] and word-final [CVʔ#] sequences is equivocal as to the phonological or phonetic character of syncope: either perspective can adequately account for this apparent phonotactic restriction.

4.6. Antigemination. Vowel deletion in Uspanteko does not normally apply between identical consonants—that is, it obeys a condition on antigemination, as seen in 12 (Hayes 1986, McCarthy 1986, Odden 1988, Baković 2005). This is a third property of Uspanteko syncope that is shared with vowel deletion in Sipakapense (Barrett 1999, 2011, Bennett 2016).

(12) Antigemination blocks vowel deletion
a. susuun  [su.'sun] ~ *[s.'sun]  ‘species of snail’
b. k’isis  [k’i.sis] ~ *[k’i.ss]  ‘cypress’
c. ájjj  [ʔá.χχ] ~ *[ʔá.χχ]  ‘sugarcane’
d. tz’únun  [ts’ún.nun] ~ *[ts’ún.nn]  ‘hummingbird’

If phonotactic sensitivity is characteristic of phonological rather than phonetic processes, this antigemination effect suggests that vowel deletion in Uspanteko is phonologically controlled. This can be compared with, for example, variable fast-speech schwa deletion in English, which routinely creates phonotactically illicit clusters

---

13 This analysis predicts that syncope might also fail to apply to unstressed vowels in […CVʔVC] and […CVʔVC] strings, which would also involve overlap with [ʔ]. This prediction is somewhat hard to test, as historical [VʔV] sequences have mostly developed into long vowels in Uspanteko, at least in roots (Campbell 1977).
(e.g. [bn] in b(a)nana [b(ə)nænə]), and which Davidson (2006b) argues is essentially phonetic in nature (see also Fougeron & Steriade 1997, Barnes & Kavitskaya 2002, Côté & Morrison 2007, Bürki et al. 2011, and related work on French schwa deletion).

However, derived geminates do occur in Uspanteko as the result of morphological affixation, as in 13. The antigemination effect observed in vowel deletion must therefore be understood as a viable constraint, if it indeed reflects a phonotactic restriction (Prince & Smolensky 2004 [1993]; see also McCarthy 1986).

(13) Geminates derived from morphological affixation

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ttaw</td>
<td>[t-taw]</td>
<td>‘it arrived’</td>
</tr>
<tr>
<td>b. xxular</td>
<td>[ʃ-ʃul-ar]</td>
<td>‘it went down’</td>
</tr>
<tr>
<td>c. jja’</td>
<td>[χ-χaʔ]</td>
<td>‘its water’</td>
</tr>
<tr>
<td>d. xojjchap</td>
<td>[ʃ-oχ-χ-ʃap]</td>
<td>‘it grabbed us’</td>
</tr>
</tbody>
</table>

A PHONETIC ACCOUNT OF ANTIGEMINATION. We now explore an alternative, phonetically oriented account of antigemination in Uspanteko. Our starting point is the assumption that vowel ‘deletion’ involves extreme gestural overlap between a vowel and its neighboring consonants, which masks the articulation of the vowel. This predicts that deletion should be less likely in contexts where vowels are phonetically lengthened, because such vowels are less likely to be entirely overlapped by adjacent consonants (§4.2).

Walter (2007) reports that vowels are phonetically longer between identical consonants in English. She proposes that vowel lengthening between identical consonants owes to universal biomechanical difficulties associated with repeating the same articulation multiple times in close succession. Walter further suggests that vowel lengthening in this context is the historical source of antigemination effects in vowel deletion cross-linguistically (see also Odden 1988:470).

Assuming that vowel lengthening between identical consonants also occurs in Uspanteko, such lengthening should lead to decreased C-V overlap in weak positions, thereby inhibiting syncope. Consequently, antigemination can be understood in purely phonetic terms, without any reference to a phonological ban on adjacent identical consonants. (We thank Jason Shaw for raising this point.)

There are two advantages to this account of antigemination. First, it is straightforwardly consistent with the fact that adjacent identical consonants do occur in Uspanteko as the result of affixation (Figure 10).

![Figure 10. [t-t] cluster tto‘n na [t-tjoʔn=na] ‘it still hurts’ (speaker 10, 2016).](image)

Such clusters may of course be consistent with a morphologically restricted version of the OBLIGATORY CONTOUR PRINCIPLE (OCP; McCarthy 1986). But if the antigemination effect in Uspanteko is grounded in vowel duration rather than a phonotactic restriction as such, then no special stipulations are needed to rule-in clusters of identical consonants produced by morphological processes.
Second, this phonetically oriented account of antigemination is consistent with the fact that some speakers do sporadically apply syncope between identical consonants, though only very rarely. This is illustrated in Figure 11. This token was produced in slow, careful speech, so deletion in this example cannot be explained away as a fast-speech reduction of some kind. It is a bona fide example of foot-sensitive, posttonic syncope.

![Figure 11. Vowel deletion in an antigemination context: áijj [ˈáj̞ɛj̞] ~ [ˈʔáχχ] ‘sugarcane’.](image)

Most speakers do not apply deletion between identical consonants, and even those few speakers who do syncopate vowels in this context do so only seldomly. Still, tokens like that in Fig. 11 are sometimes observed during formal elicitation, as well as in spontaneous speech. Antigemination effects in Uspanteko are thus near-categorical, but not truly absolute.

We propose that some speakers may simply find it easier to produce two identical consonants in close succession, lengthening the intervening vowels to a lesser degree and syncopating those vowels more often as a result. This possibility is supported by research that finds substantial individual differences in learning and executing motoric skills, including speech articulations (Byun et al. 2016, Byun & Tessier 2016, Herszage et al. 2020, Johnson et al. 2020, Anderson et al. 2021).

It would of course be possible to account for exceptions like Fig. 11 in phonological terms (e.g. Boersma & Hayes 2001:§4.4, Zuraw 2010, 2016). Still, we believe that a phonetically oriented view of antigemination has the advantage of explaining why syncope is sensitive to this phonotactic restriction specifically, and not to others, as we will see in more detail in the next section (§4.7).

We conclude that antigemination does not in fact support treating syncope as a phonological process, and may instead support a more surface-oriented analysis of syncope based on gestural overlap between vowels and consonants.

### 4.7. Deletion produces marked consonant clusters
Vowel deletion in Uspanteko produces a diverse range of consonant clusters, many of which are surprising from both a crosslinguistic and language-internal perspective. Some examples are shown in 14 and illustrated in Figure 12 (see also 20 and other examples throughout the article).

(14) Marked consonant clusters produced by deletion
- a. xíkin [ˈʃíkin] ~ [ˈʃíkn] ‘ear’
- b. íwir [ˈʔíwir] ~ [ˈʔíwr] ‘yesterday’
- c. inwach [ˈʔínwatʃ] ~ [ˈʔínwэтʃ] ‘my face’
- d. kúmb’al [ˈkúmɓal] ~ [ˈkúmɓl] ‘medicine’
- e. músmul [ˈmúsmul] ~ [ˈmúsml] ‘light rain’
- f. chukuy [tʃuˈkuj] ~ [tʃ’kuj] ‘pine cone’

### 14
The final ‘(v)’ in Fig. 11 is an apparently epenthetic vowel that sometimes cooccurs with deletion; see §4.9 below.
Figure 12. Consonant clusters produced by vowel deletion. Left: *íŋqʻab’* [ˈʔínqˀaɓ] ~ [ˈʔínqʼɓ] ‘my hand’ (speaker 15, 2016); right, *richooch* [ɾiˈʃʊoʃ] ~ [ɾˈʃʊoʃ] ‘his/her house’ (speaker 2, 2018).

Many clusters derived by vowel deletion flagrantly violate well-established phonotactic principles related to sonority, featural agreement, and so on (e.g. Parker 2002, Kehrein & Golston 2004). Depending on how these clusters are syllabified, they also violate constraints against complex onsets and codas (e.g. [ˈʃíkn]σ), constraints against unparsed, extrasyllabic consonants (e.g. [ˈʃík]σ⟨n⟩), or constraints against syllables lacking vowels (e.g. [ˈʃík.n]σ). These clusters are also aberrant from a language-internal perspective: consonant clusters are very uncommon morpheme-internally in Uspanteko (Can Pixabaj 2007:50–52), except when they arise as the result of vowel deletion; and even morphologically derived clusters tend to be restricted to the prefixal field (see examples in 8). These facts suggest that vowel deletion in Uspanteko has, at most, a very limited sensitivity to phonotactic restrictions—a characteristic that is more typical of phonetic than phonological processes (Davidson 2006b).


(15) Vowel height affects syncope
a. /i/: wálib’ [ˈwá.liɓ] ~ [ˈwá.l6] ‘my sister-in-law’
b. /u/: tukuur [tu.ˈkur] ~ [t.ˈkur] ‘owl’
c. /a/: sáq’aj [ˈsá.qʼaχ] ~ [ˈsá.qʼχ] ‘summer’
d. /e/: íntem [ˈʔín.tem] ~ *[ˈʔín.tm] ‘my chair’
e. /o/: joron [χo.ˈron] ~ *[χʼ.ron] ‘cold’

Unstressed short /a/ is highly centralized in Uspanteko and better transcribed as [ə] or [ɾ] (Bennett & Henderson 2013, Bennett, Harvey, et al. 2022). Vowel deletion can then be characterized as targeting the low-sonority unstressed vowels [i u ə/ɾ], while sparing the higher-sonority mid vowels [e o] (Crosswhite 2001, Gouskova 2003). Assuming that vowel sonority is a phonological property—albeit one with clear phonetic grounding (Crosswhite 2001, Parker 2002, 2011, de Lacy 2004, 2007b, Gordon 2006, Gordon et al. 2012)—sensitivity to sonority is a sign of a phonological process.

A phonetic account of vowel height restrictions. While the selective deletion of high and low vowels in Uspanteko can be understood as an effect of vowel sonority, a more phonetically oriented explanation for this pattern is also available. Mid vowels tend to be longer than high vowels and [ə/ɾ] (Lehiste 1970, Toivonen et al. 2015). We might then expect /i u a/ → [i u ə/ɾ] to be more susceptible to gestural overlap—that is, more deletable—than [e o] in Uspanteko, simply by virtue of their relatively short duration.
Empirically speaking, the peripheral mid vowels /e o/ do seem to be somewhat longer than other vowels in Uspanteko. In word-list data from nine speakers, Bennett, Harvey, et al. (2022) find that unstressed short /e o/ (mean duration: 81 ms, n = 151) are slightly longer than unstressed short /i u/ (mean: 70 ms, n = 368), and about as long as unstressed short /a/ (mean: 79 ms, n = 230). Stressed short vowels showed comparable durations (~105 ms) across all three vowel heights. In a separate study with twelve speakers, Bennett, Henderson, and Harvey (2022) report that mid vowels in Uspanteko are overall about 30 ms longer than non-mid vowels, though this figure pools together long and short vowels, as well as stressed and unstressed vowels.

Reference to phonetic duration also sheds light on some apparent exceptions to the generalization that mid vowels resist deletion. Vowel deletion does sometimes seem to target underlying mid vowels, as in 16 and Figure 13. The deletion of mid vowels shows a degree of lexical conditioning: deletion is common in kolob’ [k(o).ˈloɓ] ‘rope; lasso’, for example, but not in k’oxob’ [kˀo.ˈʃoɓ] ‘annatto’. Other examples of mid vowel deletion can be seen in Figs. 1 and 6 above.

(16) Deletion of mid vowels

a. ék’el [ʔék’el] ~ [ʔék’il] ‘child’
b. étzel [ʔɛtˈsɛl] ~ [ʔɛtˈsɛl] ‘bad, evil’
c. rixkólob’ [rɪʃˈkólɓ] ~ [rɪʃˈkólɓ] ‘his/her intestines’
d. kolob’ [kˈloɓ] ~ [kˈloɓ] ‘rope; lasso’

As with the avoidance of derived geminates, sensitivity to vowel height constitutes a strong trend, but not an exceptionless generalization about the conditioning of vowel deletion in Uspanteko.

It could be the case that vowel height effects are part of the categorical phonology, but are less strict for some morphemes than others. This would give rise to an overall pattern of sensitivity to vowel height, with pockets of systematic, lexically conditioned exceptions (Zuraw 2010, 2016, Coetzee & Pater 2011, Gouskova 2012, Coetzee & Kawahara 2013, Zymet 2018, Hout 2020). The intermittent deletion of mid vowels in Uspanteko could be modeled as a phonological process of this type.

15 It should be noted that individual morphemes are often produced with different vowel qualities across speakers. It is especially common for short /a e o/ to be interchanged. For example, some speakers pronounce cháqej [ʃtʃáq(e)χ] ‘dry’ (Fig. 13) as [ʃtʃáqaχ] or [ʃtʃéq(e)χ]. This is a pervasive phenomenon, and similar cases are easy to find: for example, qaleen [qaˈleːn] ~ [qeˈleːn] ‘things’, pach [ˈpaʃ] ~ [ˈpaʃ] ‘friend’, and xk ayaaj [ʃkˀa.ˈjaχ] ~ [ʃkˀa.ˈjeχ] ‘(s)he sold it’. This instability in the quality of short vowels makes it difficult to determine whether syncope in a form like [ʃtʃáqχ] really involves the deletion of a mid vowel, or instead the deletion of an underlying low vowel. Nonetheless, there are scattered cases of syncope targeting vowels that are uniformly realized as mid across speakers, such as jîleb’ [ɣtɛlɛɓ] ‘his/her shoulder’ (Fig. 13).
But as with antigemination, we believe that a phonetic perspective, grounded in physical vowel duration, may provide a deeper explanation than a purely phonological one does. In particular, a duration-based account of height-sensitive syncope in Uspanteko makes a specific prediction about where exceptions should be found: syncope should target mid vowels more often in contexts where phonetic vowel shortening is independently expected. This includes highly frequent or predictable words (Aylett & Turk 2004, Bell et al. 2009), longer words (Turk & Shattuck-Hufnagel 2000, White & Turk 2010), and faster speech rates. While other factors affect vowel duration, such as neighborhood density and contextual givenness (Aylett & Turk 2004, Vitevitch & Luce 2016), we focus here on frequency and word length because these are the two factors we can speak most confidently about in the case of Uspanteko.

We have already seen in 16 that there are some forms in which vowel deletion appears to target a mid vowel. We say ‘appears’ because in many cases it is not entirely clear whether the underlying vowel was mid or not: short vowels often vary in quality across speakers, so that, for example, chik’oor [ʃi.k’oːr] ‘hiccup’ is produced with a mid unstressed vowel [ʃo.k’oːr] by some speakers (see n. 15). The examples in 17 represent most of the forms we have encountered that show deletion of an unstressed vowel which we can confidently identify as a mid vowel, at least for some speakers (other speakers either do not syncopate in these forms, or instead use variant forms that clearly have an underlying non-mid vowel in the unstressed syllable).\footnote{In all of the forms in 17 the stressed vowel has the same quality as the syncopated unstressed vowel. We do not know if this is significant.}

\begin{align}
\text{(17)} & \quad \text{a. qeleen } [q’ele:n] \sim [q’le:n] \quad \text{‘thing; things’ } (132; 180) \\
& \quad \text{b. ójor } [’ʔóχor] \sim [’ʔóχr] \quad \text{‘(long) ago’ } (360) \\
& \quad \text{c. ék’el } [’ʔék’el] \sim [’ʔék’l] \quad \text{‘child’ } (567) \\
& \quad \text{d. étzel } [’ʔética] \sim [’ʔética] \quad \text{‘bad; evil’ } (212; 977) \\
& \quad \text{e. lékej } [’lékej] \sim [’lékχ] \quad \text{‘up; above’ } (51; 1,338) \\
& \quad \text{f. k’eten } [k’e’ten] \sim [k’e’ten] \quad \text{‘hot; sweat’ } (501; 2,474) \\
& \quad \text{g. téleb’ } [’télxb] \sim [’télb] \quad \text{‘shoulder’ } (2,148) \\
& \quad \text{h. kolob’ } [ko’loɓ] \sim [k’loɓ] \quad \text{‘rope; lasso’ } (2,418; 18,652) \\
& \quad \text{i. ixkólob’ } [iʃk’óloɓ] \sim [iʃ’kóloɓ] \quad \text{‘intestines’ } (12,957)
\end{align}

Impressionistically, these are all frequent words in Uspanteko, particularly 17a–f. Though we do not have good word-frequency estimates for Uspanteko, we can perhaps approximate those estimates by considering the frequencies of the corresponding English glosses (as obviously imperfect as this comparison is; e.g. Tang & Bennett 2018). We have noted the rank frequency of each English gloss word in 17 in parentheses, on the basis of the SUBTLEXus corpus, which contains \( \approx 60,000 \) word types and \( \approx 51 \) million word tokens for American English (Brysbaert & New 2009; \url{https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus}). Though this is a very rough and indirect measure of word frequency in Uspanteko, most of the glosses in 17 are within the top 1,000 most frequent words in SUBTLEXus. Téleb’ ‘shoulder’ and kolob’ ‘rope; lasso’ have more intermediate, though still high, frequencies. Ixkólob’ ‘intestines’ is an outlier here, though it is morphologically related to kolob’ ‘rope; lasso’, which may perhaps contribute to deletion in this form.

By contrast, words in which unstressed mid vowels clearly resist deletion plausibly have lower frequencies than the forms in 17. A sample of such forms is given in 18.
Two words, 18f,g, buck this trend by having lexical frequencies that are closer to those of the deleting words in 17.

(18) a. mesob’ [me.ˈsøb] ‘broom’ (7,095)
    b. k’oxool [kʼo.ˈʃoɬ] ‘elf (Spanish duende)’ (‘elf’: 8,184; ‘gnome’: 22,089)
    c. keqiix [kʼe.ˈʃiɬ] ‘guacamaya (species of mushroom)’
       (‘mushroom’: 11,426)
    d. wíxk’eq [ˈwʃ.kʼeq] ‘my fingernail’ (‘fingernails’: 7,743; ‘fingernail’: 14,928)
    e. k’oxob’ [kʼo.ˈʃoɬ] ‘achiote (annatto)’ (‘annatto’ does not occur in SUBTLEXus)
    f. joron [χo.ˈron] ‘cold’ (634)
    g. intem [ˈʔin.tem] ‘my chair’ (‘chair’: 1,331)

We have elicited the words in 18 many times in our own fieldwork, and we are quite confident that deletion is inhibited in these examples. We conclude that it is at least plausible that mid vowels are more likely to delete in frequent words in Uspanteko, owing to the durational and gestural reduction associated with high-frequency forms.

The second context where we have observed the otherwise unexpected deletion of mid vowels is in longer verbs like those in 19.

(19) a. xchomorsaaj [ʃt͡ʃom ˈsaɬ] ~ [ʃt͡ʃom ˈsaɬ] ‘(s)he thought it’
    b. tiq’ojomaan [tiχqˀoχ ˈmaɬ] ~ [tiχqˀoχ ˈmaɬ] ‘(s)he plays the marimba’

It has often been observed that vowels and syllables tend to be reduced in long words relative to short words (Menzerath & de Oleza 1928, Lehiste 1972, and many others). This phenomenon is sometimes known as polysyllabic shortening. Due to the morphological structure of Mayan languages, long words tend to be inflected verbs rather than nouns or other word classes (Bennett 2016, Coon 2016, Polian 2017, Tang & Bennett 2018). The prediction, then, is that mid vowels should be shortened, and thus more liable to undergo syncope, in long words, which are likely to be verbs (19). Again, our impression as fieldworkers is that this prediction is borne out.

We conclude that sensitivity to vowel height provides evidence that syncope in Uspanteko is a surface-phonetic process involving gestural overlap between consonants and shortened, reduced vowels. Appeal to phonological sonority is unnecessary and insufficient, once the phonetic duration of vowels is taken into account. Apparent lexical exceptions to height-based generalizations can be reduced to external factors, like lexical frequency, that are independently known to affect the physical phonetic duration of vowels.

4.9. Syncope and Epenthesis. Vowel deletion sometimes renders tone placement opaque. High tone cannot normally occur on the final mora—this is the condition that triggers stress retraction in tonal words with final short vowels, such as [ˈʔim.piʃ] ‘my tomato’, *[ʔim.ˈpiʃ] (§3). But when vowel deletion targets posttonic short vowels, high tone appears to be stranded on the last mora of the word, *[ʔim.pʃ].

17 Though compounding is a frequent source of longer nouns in Mayan languages (Polian 2017), it seems less common in Uspanteko than in related languages like Kaqchikel (see Can Pixabaj 2007 for examples).
18 The opaque interaction between deletion and tone placement is suggestive: Kawahara (2001, 2002, 2015) observes that many cases of opacity involve variable or optional rules, including rules like fast-speech vowel deletion that are good candidates for phonetic rather than phonological processes (see also Sanders 2003, Davidson 2006b). However, since there are cases of opacity that seem firmly phonological in nature (see e.g.
In exactly this context, a separate process of variable, word-final vowel epenthesis often applies, as in 20. The quality of the epenthetic vowel is typically \([ə]-\) like, though it can sometimes manifest as a reduced copy of the deleted vowel (cf. Can Pixabaj’s 2007:45–46 mention of ‘metathesis’).

(20) Word-final vowel epenthesis

a. kútz’ij \([ˈkútʃ̩j] \sim [ˈkútʃ̩ˈχ])\ ‘flower’
b. k’áyb’al \([k’ájɓl] \sim [k’ájɓla])\ ‘market’
c. ínchik \([ˈʔínt͡ʃk] \sim [ˈʔínt͡ʃˈk])\ ‘my work’
d. íxim \([ˈʔíʃm] \sim [ˈʔíʃm])\ ‘corn’
e. wálib’ \([ˈwálɓ] \sim [ˈwálɓ])\ ‘my sister-in-law’
f. qálaq \([ˈqálq] \sim [ˈqálq])\ ‘our plate’

An illustrative waveform is provided in Figure 14 (cf. Fig. 3 for productions of this item without epenthesis). We annotate the variable epenthetic vowel as ‘(v)’ because its quality is somewhat unstable and context-dependent.

![Waveform](image)

**Figure 14.** *inpix* \([ˈʔim.pʃ])\ ‘my tomato’, with final epenthesis (speaker 6, 2018).

The only factor conditioning epenthesis is whether deletion has applied. The composition of the word-final consonant cluster appears to be irrelevant: epenthesis occurs after voiced and voiceless consonants alike, and after clusters like 20e,f that obey sonority sequencing.

There are two possible phonological motivations for epenthesis. First, epenthesis may be a response to the opaque character of tone placement following posttonic vowel deletion (20). Insertion of a word-final vowel guarantees that high tone will occur in its normal position on the penultimate mora, resolving the opacity issue created by deletion in a somewhat roundabout fashion. Alternatively, epenthesis may serve to break up word-final consonant clusters, which typically arise only as a result of vowel deletion (e.g. Can Pixabaj 2007:50–52). In either case, epenthesis would seem to be triggered by phonological requirements, implying that it is a phonological process. But if vowel deletion precedes and feeds epenthesis, the implication is that vowel deletion must be a phonological process as well.

**Overapplication of epenthesis.** The dependence of epenthesis on a prior application of deletion is quite firm, but it too has scattered exceptions. Specifically, in words with penultimate accent, epenthesis can sometimes be observed even in the absence of deletion (Figure 15; cf. Fig. 4). In such examples, the nondeleted vowel is always reduced.

---

Vowel deletion as grammatically controlled gestural overlap

The fact that epenthesis can be partially dissociated from deletion thus undermines the use of epenthesis as a diagnostic for the phonological status of vowel deletion in Uspanteko.

In §8.4 we develop an explicit analysis of reduction-dependent epenthesis in terms of gestural overlap: essentially, the word-final consonant overlaps the preceding vowel to such an extreme extent that the vowel ‘peeks out’ on the other side of the consonant, as in 21 (Steriade 1990, Blevins & Garrett 1998, 2004, Yanagawa 2003).

(21) Schematic pattern of articulatory overlap resulting in ‘epenthesis’

If ‘epenthesis’ is simply another instance of articulatory overlap, then apparent cases of overapplication (Fig. 15) are no longer quite so mysterious: like regular cases of epenthesis, they reflect a particularly extreme pattern of gestural overlap and retiming between vowels and consonants. We defer a more detailed explanation of this proposal until §8.4, after outlining the basics of our formal analysis.

4.10. Sensitivity to foot structure. We argued above that vowel deletion is sensitive to metrical foot structure in Uspanteko. The key evidence supporting this claim is the selective nature of vowel deletion (§3.1). Vowel deletion aggressively targets unstressed vowels that can be construed as foot-internal: pretonic vowels in words with final stress, […σσ] (22), and posttonic vowels in words with penultimate stress, […σσσ] (23). Unstressed vowels are much less likely to delete in other positions, as is clear when considering relatively long words like those in 22–23. (morphological effects on syncope have been controlled in these examples; see §4.4.)

(22) Pretonic deletion
   a. xqaq’asaj [ʃqaq-reduxaʃ] ~ [ʃqaq-reduxaʃ] ‘we passed it’
   b. tijq’ojomaan [tiʃq-reduxoʃmaʃ] ~ [tiʃq-reduxoʃmaʃ] ‘(s)he plays the marimba’

(23) Posttonic deletion
   a. alib’xlēb’ [ʔaliɓ-reduxʃiɓ] ~ [ʔaliɓ-reduxʃiɓ] ‘daughters-in-law’
   b. chapālik [ʃa-rexapalik] ~ [ʃa-rexapalik] ‘seized’

Metrical foot structure is without doubt abstract and phonological in nature. If Bennett and Henderson (2013) are correct that the position of vowel deletion is conditioned by foot structure, this counts as a clear argument that vowel deletion is grammatically controlled in Uspanteko, be it a case of phonological deletion or language-specific phonetic patterning.

4.11. Pre- vs. posttonic deletion. We have shown that the locus of syncope in Uspanteko is dependent on the position of stress: when stress is final, syncope targets
the pretonic position [...σσ], and when stress is penultimate, syncope targets the posttonic position [...σσ] (§3.1). Furthermore, syncope is variable in both positions. However, it has long been evident to us in our fieldwork that syncope is not equally common in these two positions. Posttonic syncope [...σσ] is quite common, and pretonic syncope [...σσ] clearly less so. Why would this be?

One possibility, suggested to us by Marc Garellek, is that unstressed vowels may be phonetically longer in pretonic than in posttonic position. If longer vowels are less susceptible to gestural overlap, then phonetic lengthening in pretonic position could derive the observation that syncope is less common in pretonic syllables (see also §4.2, §4.6, and §4.8 above).

This suspicion seems to be on the right track. In word-list data from nine speakers, Bennett, Harvey, et al. (2022) found that stressed short vowels averaged about 105 ms (n = 795). Unstressed short vowels averaged 75 ms in pretonic position (n = 717), but only 60 ms in posttonic position (n = 32; p < 0.05 via two-sided t-test; see also §4.8). Pretonic and posttonic vowels have very different rates of attestation in this data set, in part because many pretonic vowels were in prefixes or word-initial position, two contexts in which syncope is absolutely prohibited (§4.4, §4.5). Syncope is also more common in posttonic position, so there were simply fewer audible vowels to measure in posttonic syllables. Still, assuming that these figures are representative of Uspanteko as a whole, then durational differences between pretonic and posttonic unstressed syllables might account for the greater prevalence of posttonic syncope.

It remains to be explained why pretonic lengthening might occur in the first place in Uspanteko. In any case, Uspanteko is not the only language reported to show pretonic vowel lengthening: similar phenomena have been observed in at least Tiberian Hebrew (McCarthy 1981, Malone 1990, Churchyard 1999), Canadian French (Walker 1984:Ch. 3), Córdoba Argentinian Spanish (Lang-Rigal 2014), and various Slavic languages (Borise 2017). And in the realm of deletion, LoCasto and Connine (2002) and Patterson et al. (2003) report that [ə]-deletion in English—which arguably reflects gestural overlap rather than true deletion (Davidson 2006b)—is less common in pretonic than posttonic position, consistent with our observations for Uspanteko. Similarly, Caballero (2008:Ch.2) reports that vowels in Choguita Raramúri are more likely to reduce and delete in posttonic than in pretonic position, and Oh (2021) provides analogous results for vowel reduction in Brazilian Portuguese (see also Crosswhite 2001).

Once again, appeal to surface phonetic factors—particularly vowel duration—provides insight into the fine details of Uspanteko syncope. This explanatory gain provides additional support for analyzing vowel deletion in Uspanteko in terms of gestural overlap between vowels and consonants.

5. INTERIM SUMMARY. We have now surveyed a range of diagnostics to assess (i) whether syncope in Uspanteko is under grammatical control, and if so, (ii) whether the relevant notion of ‘control’ implicates phonological vowel deletion or a language-specific phonetic pattern of high gestural overlap between vowels and flanking consonants.

Some of these diagnostics clearly indicate that vowel deletion in Uspanteko is grammatically controlled. Deletion is sensitive to morphological structure and metrical footing, and does not depend on speech rate or style. These are characteristics commonly associated with phonological processes; at a minimum, they indicate that deletion is controlled and planned by speakers, rather than being a mechanical by-product of speaking quickly. These diagnostics are thus consistent with treating ‘deletion’ as a controlled, language-specific pattern of extreme gestural overlap in [CVC] sequences.
Compelling evidence for an articulatory treatment of vowel deletion comes from the close relationship between deletion and vowel reduction. Vowel reduction occurs in exactly the same contexts as deletion in Uspanteko. This suggests that deletion may be the extreme endpoint of a gradient, continuous pattern of vowel reduction in weak positions, involving increasing amounts of gestural overlap between prosodically weak vowels and their neighboring consonants.

We also considered whether syncope shows sensitivity to phonotactic restrictions. The answer is a very qualified ‘yes’: deletion does not occur in [#ʔVC] and [CVʔ#] sequences, is inhibited between identical consonants, and primarily targets short /i u a/. But we have argued that these apparent phonotactic restrictions can also be attributed to phonetic factors. First, gestural overlap with [ʔ] usually results in creakiness, which does not acoustically mask vowels. Second, vowel deletion should be less likely for phonetically longer vowels, and more likely for shorter vowels. Both antigemination and the resistance of /e o/ to deletion may then stem from contextual phonetic vowel lengthening, which is independently attested in each of these contexts. Sporadic exceptions to these patterns can be similarly attributed to pressures favoring vowel shortening (e.g. lexical frequency). On this view, syncope shows no direct phonotactic conditioning at all—this is consistent with the fact that vowel deletion often derives marked and otherwise unattested consonant clusters (e.g. inchikich [ʔinˈtʃíktʃ] ~ [ʔinˈtʃiˈkʃ] ‘my basket’).

Lastly, deletion appears to feed a process of word-final vowel epenthesis. However, as outlined in §4.9, even ‘epenthesis’ can be construed as a consequence of extreme gestural sliding, an intuition we develop in more detail in §8.4.

In the next section we present instrumental phonetic evidence that supports the claim that vowel ‘deletion’ in Uspanteko reflects high levels of gestural overlap between vowels and flanking consonants. Weak vocal-fold vibration associated with ‘deleted’ vowels can sometimes be detected, via electroglottography, for vowels that appear to be entirely absent from the acoustic signal. This implies that even ‘deleted’ vowels are present in the input to articulatory planning and are not deleted in the phonology proper.

In §7 we provide a formal analysis of vowel deletion in terms of grammatically controlled gestural overlap, following work in **ARTICULATORY PHONOLOGY** (e.g. Gafos 2002).

6. A pilot EGG study on vowel deletion. If vowel deletion in Uspanteko is categorical and symbolic, there should be no trace of the deleted vowel in the surface phonetics. If, by contrast, ‘deletion’ involves high levels of gestural overlap, then some phonetic signature of the overlapped vowel might be present in the phonetic form of the utterance.

To explore this prediction we carried out a pilot study on the phonetics of vowel deletion in Uspanteko, using audio recordings collected with simultaneous electroglottography recordings. Electroglottography (EGG) is a noninvasive technique for measuring vocal-fold contact during speech; for references and overviews, see Garellek 2013, Herbst 2020, and Kochetov 2020.

Here, we use EGG to detect the presence or absence of voicing during consonant clusters derived by vowel deletion. Our research question is the following: is there electroglottographic evidence that vocal-fold contact associated with deleted vowels is (partially) retained, even when the vowel is not apparent in the acoustic signal? If so, this would support our hypothesis that ‘deletion’ involves high degrees of gestural overlap rather than the categorical elimination of a segment in the symbolic phonology.
The signature of such ‘covert’ vowels would be (i) no clear evidence for a vowel in the acoustic recording, but (ii) an oscillating signal in the EGG recording corresponding to (weak) voicing that can be attributed only to an underlying vowel, and not to the flanking consonants (e.g. /...qaχ#/ → [...qχ#]: Figure 16). The voicing during such intervals might be attenuated, but the mere fact of voicing would provide evidence that even ‘deleted’ vowels are sometimes retained in the surface phonetics, even when not readily apparent in the acoustic signal (see also Gick et al. 2012).

Covert voicing should be possible under several aerodynamic conditions involving weak, attenuated vocal-fold vibration. Weak voicing involves relatively low-amplitude vocal-fold oscillation and a relatively low fundamental frequency (Pinho et al. 2012). During stop closure, voicing is frequently weak for aerodynamic reasons (e.g. Burton et al. 1992, Solé 2018). Additionally, when voicing is weak, the acoustic energy associated with voicing may be mostly or entirely absorbed by soft tissues in the oral cavity (Johnson 2012:Chs. 8.1.2, 9.1, etc.). In such cases, no audible noise associated with voicing should radiate out from the vocal tract. The result is vocal-fold vibration without audible consequences—covert voicing.

Figure 17 illustrates how a brief period of covert voicing can occur during voicing coarticulation: weak carryover voicing on [ɓ] from the preceding vowel is visible for several periods in the EGG signal (bottom panel) after essentially disappearing from the audio signal (top panel) (implosives otherwise tend to be voiceless in K’ichean languages; Bennett 2016, Bennett, Harvey, et al. 2022). For comparable EGG recordings from other languages, see Mazaudon & Michaud 2008:Fig. 6, DiCanio 2012:Fig. 12, and Recasens & Mira 2012:Fig. 4.

In fricatives, aerodynamic constraints also lead to weak voicing during closure (Ohala 1983, 1993, Pinho et al. 2012). Since fricatives are produced with continuous oral airflow, periodicity can normally be detected in the audio signal during a fricative. However, we hypothesize that in cases of exceptionally weak voicing, the same damping
effect seen in stops (Fig. 17) may render periodic voicing too weak to hear or identify in the acoustic signal. Furthermore, fricative noise itself may interfere with the perception of weak voicing. For example, the dorsal fricatives [x ɣ] have a low-frequency periodic component in the range of f0, which reflects slow vibration of the uvula during their articulation (Redmon & Jongman 2018). This low-frequency noise may partially mask the vibration of the vocal folds, as an instance of destructive interference. (We thank John Kingston for discussion of this point.)

In summary, weak, covert voicing is an empirical possibility during phonologically voiceless obstruents. If such covert voicing is observed in clusters of voiceless obstruents derived by syncope, /C̥VC̥# → [C̥C̥#], that voicing must be a phonetic vestige of the underlying vowel. Such a result would indicate that ‘syncope’ does not in fact involve the categorical elimination of a vowel from phonological representations.

We primarily report on EGG and acoustic data from one speaker here, though we also consider more tentative findings for three other speakers.

6.1. Recording. Participants read a list of words in Uspanteko, which were presented in isolation on a laptop screen using a custom Python script. Presentations were self-paced: participants read each item twice before proceeding to the next word by pressing the space bar.

Audio was recorded using a head-mounted microphone (Audio-Technica ATM73a) and solid-state portable recorder (Zoom H5), at a 48 kHz sampling rate with 24 bit quantization. The EGG signal was recorded using a Glottal Enterprises model EG2-PCX2 electroglotograph. The EGG signal was recorded with a 44.1 kHz sampling rate, and high-pass filtered with a 10 Hz threshold. The recordings were made in a quiet room in San Miguel Uspantán in summer 2019. Recording sessions lasted thirty to forty-five minutes.

The audio recordings were synced with the EGG recordings by aligning the first glottal pulse in the audio signal with the first glottal pulse in the EGG signal, across seven vowels, to determine the lag between the two recordings. The timecode of the EGG recording was shifted by the mean of these lag measurements to achieve alignment.

Seven speakers of Uspanteko were recorded in this study, but technical issues with the EGG machine impacted data quality for all but one of them. In this section we report our full findings for speaker 6, as well as some more limited findings for three other speakers whose data was partially usable. The entire recording session for speaker 6, including both EGG and audio recording, is available as supplemental material online at http://muse.jhu.edu/resolve/200.

Materials. The word list used for this study included fifty-one items. Most of these items were included in the list in order to study voice quality on tonal and nontonal vowels, which was the main purpose of data collection for these sessions, rather than the study of vowel deletion as such (we do not discuss voice quality at all here). Each speaker read through the entire word list five times.

A number of items in the word list had the potential for posttonic syncope, for example, inchik [ʔin.ʃik] ‘my work’. Whether syncope actually applied to any given form was (i) dependent on each speaker’s vocabulary (e.g. some speakers have inchaak [ʔin.ʃak] instead of inchik [ʔin.ʃik] for ‘my work’), and (ii) variable across productions, as is typical of vowel deletion in Uspanteko (§3.1). We limited the analysis of deletion to word-final /CVÇ#/ sequences in which both flanking consonants were voiceless. This ensured that any voicing observed in tokens of deletion /CVÇ# → [ÇÇ#] could be attributed only to the underlying vowel. Many
of the items showing vowel deletion in our recordings could not be used for the present
analysis because at least one of the flanking consonants was voiced, for example, rixóqil
[ri·ʃó.q(i)] ‘his/her wife’. We focused on posttonic syncope simply because there were
fewer cases of pretonic syncope to examine in this data (see also §4.11).

Derived [CC#] clusters were compared to (i) tokens of the same items in which deletion
had not applied, yielding surface [CVČ#], and (ii) underlying voiceless /C̥C/# clusters.
Underlying voiceless /C̥C/ clusters provide a baseline for the absence of voicing
in the EGG signal, while surface /CVČ#/ → [CVČ#] sequences provide a baseline for
vowel-related voicing in the EGG signal. To illustrate, voicing in syncopated /ʃáq(a)x/ →
[ʃáq] ‘dry’ was compared to voicing in nonsyncopated /ʃáq(a)x/ → [ʃáq], as well as
to underlying [f-qa-kʰam] ‘we received it’ (among other comparisons). The underlying
voiceless /C̥C/ sequences analyzed here include word-initial, word-medial, and word-
final clusters.

6.2. Coding and measurement. /CVČ#/ intervals were coded as having undergone
deletion (or not) by an inspection of the waveform and spectrogram for that interval. If
vowel-like periodicity with higher formant structure was observable in the waveform
or spectrogram, the token was coded as retaining the underlying vowel (Figure 18, left-
hand side). If vowel-like periodicity was absent, the token was coded as having under-
gone vowel deletion (Figure 18, right-hand side; see Bürki et al. 2011).

![Figure 18. Sample waveforms, spectrograms, and pitch traces for variable vowel
deletion in cháqaj [ʃáq(a)x] ‘dry’ (speaker 6, 2019).](image)

For speaker 6, we analyzed 209 tokens in the underlying /CC/ condition, fourteen
tokens of apparent vowel deletion, and eighteen tokens in which a vowel was
still detectable in the audio signal. The small number of tokens for the underlying
/CVC#/ → [CC] → [CVČ#] conditions reflects the fact that only a minority of the items
in our word list met our criteria for inclusion in the analysis (recall that these recordings
were originally intended to investigate voice quality, not syncope).

There were five target items in the /CVČ#/ condition: cháqaj [ʃáq(a)x] ‘dry’, chí-
kich [ʃík(i)x] ‘basket’, lékej [lek(e)x] ‘up, above’, ichaj [ʔitf(a)x] ‘herb’, and inchaj
[ʔinf(a)x] ‘my pine cone’. Actual cases of syncope were observed for only three of
these items (chéqaj, lékej, and chíkich).

6.3. Results. As predicted by the gestural-overlap analysis, voicing was still detect-
able in the EGG signal in at least some tokens in which ‘deletion’ had taken place in the
audio signal. Syncopated tokens were classified as having clear covert voicing if there
were salient oscillations in the EGG signal; as having possible ‘weak’ covert voicing
if some periodicity was visible, but extremely low in amplitude; and as lacking covert
voicing otherwise (i.e. if the EGG signal was essentially flat). Examples of each of these
phonetic types are provided in Figure 19.
Vowel deletion as grammatically controlled gestural overlap

The vocal folds may approximate closure without vocal-fold contact appearing clearly in the EGG signal (Orlikoff et al. 2009, Herbst 2020). Consequently, even marginal cases of low-amplitude vibration in the EGG signal may indicate a glottal gesture, albeit a reduced one involving incomplete glottal adduction. For that reason, we included cases of ‘weak’ covert voicing in our analysis.

Table 2 provides a summary of our observations. As is typical in Uspanteko, the same items were sometimes produced with syncope, and sometimes without. This within-speaker variation was observed during a single recording session and under very formal elicitation conditions. When syncope did apply, covert voicing was visible in 57% of productions (8/14 tokens).

Figure 20 illustrates clear covert voicing for one token of cháqaj ['tʃáq(a)χ] ‘dry’, corresponding to the right-hand panel of Fig. 18 above. Voicing during the derived consonant cluster /…qχ#/ → […qχ#] can only be attributed to the underlying, ‘deleted’ vowel, since the flanking consonants are both voiceless. This voicing is weaker (lower amplitude and lower frequency) than the voicing observed in the stressed [ˈá] that precedes it.19

As noted in Table 2, covert voicing of the type seen in Fig. 20 was never observed in underlying clusters of voiceless consonants /C̅C̅/ (e.g. jqu [ʄ-ʄu] ‘my neck, throat’; see also Fig. 19). This rules out any alternative explanation of covert voicing that attributes such voicing to properties of the consonants themselves, alone or in combination in a cluster.

Though we are hesitant to make statistical claims based on the limited counts in Table 2, a Fisher’s exact test ($p < 0.001$) detects a significant association between cluster type (underlying vs. derived) and the presence or absence of covert voicing (collapsing together

---

19 Voicing is often assumed to be the default state of the vocal folds in articulatory phonology (Goldstein & Browman 1986, Browman & Goldstein 1992a, Chitoran et al. 2002, Smorodinsky 2002). From that perspective, the vestigial, covert voicing seen in Fig. 19 represents a gap between two glottal-spreading gestures for voicelessness, rather than a voicing gesture as such (see also Munhall & Löfqvist 1992).
cases of ‘possible’ and ‘clear’ covert voicing). This is consistent with our claim that covert voicing owes to the underlying lexical vowel in syncope contexts: covert voicing occurs in 57% of derived consonant clusters (8/14), but 0% of underlying consonant clusters (0/209), despite the fact that there are many more underlying clusters in our data. (A χ²-test with simulated p-value based on 2,000 replicates also returns a significant p < 0.001.)

Table 2 shows that covert voicing occurs in derived [qχ#] and [kχ#] clusters, while [ktʃ#] clusters show essentially no EGG evidence of ‘hidden’ voicing in tokens where the vowel is absent from the acoustic signal. This asymmetry may reflect aerodynamic factors, if voicing is harder to produce in plosive-plosive clusters than in plosive-fricative clusters (Ohala 1983, Westbury & Keating 1986, Davidson 2016, 2018).

We focus here on posttonic syncope simply because there were fewer cases of pretonic syncope to examine in this data (see also §4.11). However, at least one case of pretonic syncope in this data set appeared to show covert voicing as well (Figure 21), though it is perhaps possible to identify the spectral change at the end of the [s] as acoustic evidence of voicing with concomitant higher formant structure. Perceptually, voicing between [s] and [qʾ] in this token is inaudible to us: we hear an [sqʾ] cluster (see Bürki et al. 2011 for related discussion).
Finally, we consider some additional data from three other speakers. The EGG recordings for six of the seven speakers in our pilot study were marred by recording artifacts and/or excessive noise that precluded the use of that data for any in-depth analysis. Still, usable tokens could be extracted from the recordings for speakers 1, 5, and 7 in our pilot study, at least with some additional postprocessing (e.g. targeted band-pass and stop filtering to remove line noise, the remove noise function in PRAAT, etc.; Boersma & Weenink 2020). For all three of these speakers, tokens involving syncope can be found that clearly show voicing in the EGG signal that is absent from the corresponding audio recording, or is at least not very obvious. Some illustrative tokens are provided in Figure 22.

We conclude that covert voicing in syncopated /CVC/ sequences is unlikely to be an idiosyncrasy of the particular audio recording analyzed here for speaker 6. Instead, covert voicing is most likely a general phenomenon in syncope contexts in spoken Uspanteko, for at least some speakers.

To summarize, pilot EGG data suggests that even ‘deleted’ vowels in Uspanteko are sometimes present in the surface phonetics as attenuated voicing associated with the ‘deleted’ vowel. This is consistent with the claim that vowel deletion in Uspanteko is the result of high gestural overlap rather than the categorical deletion of a vowel target in the symbolic phonology.

7. DELETION AS PHONOLOGICALLY CONTROLLED GESTURAL OVERLAP. Vowel deletion in Uspanteko appears to be the endpoint of a gradient process of vowel reduction, involving different degrees of overlap between vowels and flanking consonants (§4.2). Electroglossographic evidence (§6) provides strong confirmation of this hypothesis, as even vowels that leave little or no trace in the acoustic record may be realized as covert, low-amplitude phonation during derived /CVC#/ → /C/C# intervals.

But this is not to imply that deletion is ‘merely’ phonetic, in the sense of being some inevitable by-product of the mechanics of speech production. To the contrary, there are
compelling reasons to believe that vowel deletion is under speaker control. Deletion is a systematic, salient characteristic of spoken Uspanteko. Even if deletion amounts to gestural overlap, such overlap represents a language-specific pattern of articulatory coordination, which must be learned (Keating 1984b, 1990a, Liberman & Pierrehumbert 1984, Pierrehumbert 1990, Kingston & Diehl 1994). Deletion shows phonological conditioning, being sensitive to foot structure and at least some segmental phonotactics. Deletion is also restricted by morphological structure, being prohibited absolutely from applying in prefixes. Lastly, speech rate and style have no discernible effect on whether deletion may apply. These are the traits of an intentional, grammatically determined process.

The force of these observations is clear: vowel deletion in Uspanteko is a grammatically controlled pattern of gestural overlap. This echoes previous proposals that treat certain gradient patterns of vowel reduction as essentially phonetic processes that are nonetheless sensitive to grammatical considerations such as syllable structure (e.g. Kondo 1994, 2008, Beckman 1996, Purse 2019).

It is worth reiterating that deletion, understood as extreme gestural overlap, regularly occurs in contexts that are antagonistic to articulatory reduction. These include slow speech and environments like utterance-final position that induce phonetic lengthening (e.g. Myers & Hansen 2007, Byrd & Krivokapić 2021, Paschen et al. 2022 and references there). The fact that deletion routinely occurs in these contexts provides further evidence that foot-internal vowel reduction and gestural overlap are under speaker control. Deletion is planned and intentional, even if only unconsciously so.

To fully specify this proposal, in the following sections we provide an explicit analysis of deletion, based on language-specific patterns of articulatory coordination, which interact with abstract phonotactic and morphological principles.


8.1. Gestural coordination and overlap. Articulatory phonology assumes that segments like [s] are composed of sets of articulatory gestures. These gestures include oral constriction goals (e.g. alveolar closure), as well as gestures regulating laryngeal state (e.g. spread vocal folds) and nasality (e.g. raised velum). The internal temporal structure of those gestures is represented in the grammar (Figure 23). The phonology itself controls coordination between gestures by specifying the extent to which gestures overlap with each other, both within individual segments (e.g. [ɪs]) and between successive segments (e.g. [np]) (Browman & Goldstein 1990, Gafos 2002, Davidson 2003, Hall 2003, Borroff 2005, 2007, Bradley 2007, Smith 2018, Walker & Proctor 2019).
In articulatory phonology, gestures are abstract: they correspond to vocal-tract goals (e.g. lip closure), rather than specific movements of the articulators. The actual trajectories of articulators can be lawfully derived from abstract gestural representations, by means of general principles of physical movement and coordination (e.g. Saltzman & Munhall 1989), but the gestural representations themselves are conceptually distinct. As we will see, this abstractness is important: it allows for representations in which two incompatible articulations are specified to occur at the same time. Such representations can be used to model coarticulation, assimilation, and other phenomena potentially related to gestural overlap (see also Browman & Goldstein 1992a, Iskarous et al. 2012, Bellik 2018, Smith 2018, and others).

In principle, gestures and segments could be specified to have any degree of overlap whatsoever, to an arbitrary level of precision. In practice, patterns of coordination are drawn from a smaller set of possible options (Browman & Goldstein 1990, Gafos 2002, Davidson 2003, Tilsen 2016, Zsiga 2021). Gestures are decomposable into landmarks, which correspond to important articulatory events like the achievement and release of a constriction (Fig. 23a). Gestures are coordinated with each other through reference to these landmarks. Segments, being composed of gestures, are sequenced and coordinated in the same way: in *mat* [mæt], for example, the tongue-tip gesture for [t] is timed to begin near the offset of the dorsal gesture for [æ] (Nam et al. 2009).

Two patterns of intersegmental coordination are particularly relevant for our analysis. The first pattern involves little or no overlap between successive segments (Fig. 23b). Here, the onset of the second segment is aligned to the offset of the first. This is also known as antiphase coordination: the gesture for the first segment completes before the gesture for the second segment is initiated, so the two gestures are out of phase (Saltzman & Munhall 1989, Goldstein et al. 2006). In this pattern of coordination, the release of the first segment should be quite audible, as the interval between the two segments involves a very open vocal tract (e.g. Gafos 2002, Smorodinsky 2002, Goldstein 2011).

The second pattern of coordination involves total or near-total overlap between adjacent segments. Here, the onsets of both segments are essentially synchronous. This is known as in-phase coordination (Fig. 23c). In the following section we lay out the phonetic consequences of articulatory overlap, which can arise under in-phase coordination as well as other articulatory regimes.

CONSEQUENCES OF GESTURAL OVERLAP. Research on gestural coordination in CV syllables has found that the lingual gesture for the vowel begins during the articulation of

---

20 Readers familiar with Gafos 2002 may notice that the shorthand notation we use for gestural coordination patterns in Fig. 23 differs from the shorthand notation Gafos uses. This is simply because we find the meanings of the notational symbols in Fig. 23 easier to remember.
the consonant itself. The vowel may begin near-simultaneously with the consonant (Fig. 23c; e.g. Goldstein et al. 2006) or may be coordinated with the center of the consonantal constriction (Browman & Goldstein 1990, Shaw et al. 2009), a possibility we set aside here. Even though the consonant and vowel gestures are highly overlapped, the vowel remains audible because vowel gestures take longer to execute: as a consequence, the slower vowel gesture persists past the end of the consonant, and the vowel can be heard (Figure 24 and §4.2; e.g. Goldstein et al. 2006, Pouplier 2011).

Figure 24. In-phase \([C \approx V]\) coordination and overlap in a simple \([CV]\) syllable.

When the consonant and the vowel involve distinct articulators, as in /bi/, the articulatory consequences of C-V overlap are limited. However, when the consonant and vowel draw on the same articulators, as in /ɡi/, overlap may lead to competition between two conflicting gestural targets (e.g. the incompatible dorsal constrictions for /ɡ/ and /i/). The typical outcome when antagonistic gestures overlap is blending: the competing gestures become similar to each other. So for example, in English /ɡi/, the /ɡ/ is substantially fronted, and the /i/ somewhat backed (e.g. Öhman 1966, Saltzman & Munhall 1989). This represents an articulatory compromise between the two competing constrictions.

Alternatively, one of the segments may be prioritized and may completely dominate the output, thereby obscuring the other segment entirely. This can occur when one of the segments is highly reduced, as in Uspanteko. Coarticulatory dominance may also occur when overlapping segments have very different blending strengths, modeled using the \(\alpha\) parameter in articulatory phonology. For reasons of space we do not discuss the \(\alpha\) parameter in any detail here, though it could easily be integrated into our analysis; see Recasens 1999, Iskarous et al. 2012, Miller 2013, Smith 2018, and Walker & Proctor 2019 for more details.

Finally, articulatory overlap between gestures can lead to the wholesale masking or hiding of entire segments. For example, Browman and Goldstein (1990) argue that apparent /t/-deletion in phrases like perfect memory \([pʰɹɛfɪk(t) \# mɛməɹi]\) actually involves articulatory overlap between [t] and the following [m], such that the lip closure for [m] hides the tongue-tip release for [t], rendering it inaudible. Overlap of this type, between consonants and reduced vowels, is central to our analysis of syncope and covert voicing in Uspanteko.

Syllable-internal coordination. We follow previous work in articulatory phonology in assuming that onsets and codas are coordinated in different ways with the syllable nucleus (Gafos 2002, Goldstein et al. 2006, Nam et al. 2009, Gafos & Goldstein 2012, Mücke et al. 2020). Onset consonants are coordinated in phase with the following vowel, while coda consonants are coordinated in an antiphase relation with the vowel (Fig. 23b, and Figure 25). Effectively, this means that onset consonants are highly overlapped with the vowel nucleus, while coda consonants are specified to overlap with the preceding vowels to a much lesser degree. Browman and Goldstein (1990) propose a similar pattern of antiphase coordination between \(V_1\) and C for intervocalic consonants \(/V_1CV_2/\).

These patterns of in-phase and antiphase coordination are also relevant for the analysis of Uspanteko syncope. Primarily, we must make sure that antiphase coordination
is sufficiently weak in syncope contexts to allow for massive gestural overlap between vowels and coda consonants in posttonic position [CV.CV.C], and between vowels and following consonants in pretonic position [CV(C).CV].

In our formal analysis of syncope, we invoke the notion of COUPLING STRENGTH (Byrd 1996, Browman & Goldstein 2000, Cho 2001, Goldstein & Fowler 2003, Bellik 2018). The basic idea is that some coordination relations are enforced more strictly than others—the gestures involved are more tightly coupled and have a more stable timing relationship. Conversely, weakly coupled gestures have less stable and more variable coordination patterns. Weakening the strength of an antiphase timing relationship in VC contexts may thus lead to greater variability and greater overlap in VC coordination (e.g. Browman & Goldstein 1990, Nam et al. 2009), increasing the likelihood of apparent vowel ‘deletion’.

8.2. PROSODIC GESTURES: EXPANSION AND REDUCTION. Finally, we consider how prosodic strength and weakness—the relative prominence of different syllables—can be modeled in a gestural framework. Articulatory gestures are longer, faster, and less overlapped in stressed positions (e.g. de Jong 1995). These effects are modeled in articulatory phonology by means of prosodic modulation gestures, called µ-gestures (Byrd & Krivokapić 2021, Krivokapić 2022 and references there). A µ-gesture is an abstract instruction to increase the spatial parameters of a gesture: gestures become more extreme or hyperarticulated, and increases in gestural duration and velocity follow as consequences of those spatial changes (e.g. Katsika 2018, Roon et al. 2021). These effects are schematized in Figure 26.

We extend these ideas by proposing that µ-gestures can also be used to generate articulatory REDUCTION in nonprominent positions. It has been suggested that prosodic modulation gestures like µ-gestures can vary in their strength (e.g. Byrd & Krivokapić 2021; see also Jun 1996). Oversimplifying the actual implementation of this suggestion, let us assume that a µ-gesture with strength equal to 1 has no effect on gestural magnitude, while a µ-gesture with strength greater than 1 will lead to some degree of

---

21 Some work in articulatory phonology distinguishes µ-gestures that modulate articulatory magnitude from µ-gestures that modulate articulatory speed (see e.g. Byrd & Krivokapić 2021). We abstract away from this distinction here.
hyperarticulation and gestural expansion. If we allow the strength of a μ-gesture to take on a value below 1, it will have the effect of shrinking the magnitude of any associated gestures and slowing the speed at which those gestures are executed. This is, of course, the hallmark of articulatory reduction and undershoot (Fig. 26).

Syncope in Uspanteko depends on prosodic prominence: weak, foot-internal vowels undergo syncope, while stressed vowels and foot-external vowels do not. We can thus model reduction in weak, foot-internal position by associating reduced syllables with a μ-gesture of strength less than 1 in this context. If values for μ are allowed to range as low as zero, it is even possible for vowels to be ‘reduced’ all the way to nothing (Ernestus 2011), such that true, categorical deletion is modeled as the literal endpoint of gradient reduction.

8.3. GESTURAL OVERLAP IN THE GRAMMAR.

VOWEL REDUCTION. In this section we implement the proposal that a ‘weakening’ μ-gesture can be used to model foot-internal reduction of unstressed syllables (§8.2). The key constraint is in 24.

(24) REDUCE-IN-FOOT: Assign one violation for every metrical foot that does not have a μ-gesture of strength < 1 associated with its weak, unstressed branch.


(25) Dep[μ-gesture]: Assign one violation for every μ-gesture present in the output, but not in the input.

Dep[μ-gesture] is analogous to constraints like Ident[stress], which penalize the addition or removal of stress prominence in output forms (Pater 2000). Dep[μ-gesture] similarly regulates the addition of both prominence-lending μ-gestures and weakening μ-gestures with strength < 1.

Since vowels are produced in phase with preceding onset consonants, such that their gestures initiate at roughly the same time, reducing the magnitude and the duration of a vowel will lead to greater proportional overlap between that vowel and the preceding

22 Nam et al. (2008) and Saltzman et al. (2008) develop an alternative approach to prosodically conditioned reduction. They propose that speech planning is influenced by a foot-level oscillator that tries to keep the duration of metrical feet constant. In languages where this oscillator is prioritized, the lengthening of stressed syllables may trigger the shortening of unstressed syllables: if stressed (σ́σ) gets long, unstressed (σσ́) must get short, in order to keep the overall duration of the foot relatively fixed (compare with the traditional notion of ‘stress-timed’ languages; e.g. Arvaniti 2012, Krivokapić 2022).

The framework developed by Nam et al. (2008) and Saltzman et al. (2008) is philosophically similar to our approach: we both model stress-related durational effects with μ-gestures, and both assume that foot structure affects segment duration. However, Nam, Saltzman, and their colleagues are interested in a different reduction phenomenon—polysyllabic shortening (§4.8)—than is our focus here. Consequently, the notion of ‘foot’ they invoke is also quite different: it corresponds to interstress intervals, of arbitrary length, that can span word boundaries (e.g. the four-syllable phrase big for a duck counts as a ‘foot’ for them; see also Keating 1988, Krivokapić 2022). This departs from the strictly binary, word-bounded (σσ) feet we assume for Uspanteko. For that reason, we leave a more detailed comparison of these two approaches for another occasion.
onset consonant (26). The subscript $\mu_R$ indicates the presence of a reducing $\mu$-gesture on a vowel or syllable.  

\[ (26) \text{REDUCE-IN-Foot} \gg \text{Dep}[\mu\text{-GESTURE}] \]

a. **FOOT-INTERNAL WEAK SYLLABLES:** audible, reduced vowel

\[
\begin{array}{c}
C \quad V_{\mu R} \\
\end{array}
\]

b. **ELSEWHERE:** audible full vowel

\[
\begin{array}{c}
C \quad V \\
\end{array}
\]

While 26 illustrates vowel reduction in posttonic position, exactly the same constraints will produce vowel reduction in pretonic position \[\ldots(CV_{\mu R}(C).\sigma)\], assuming again that both positions are foot-internal.

Vowel reduction is gradient in Uspanteko, in the sense that reduced foot-internal vowels differ in their degree of shortening and undershoot (§4.2). This variability could reflect differences in the strength of the $\mu R$ gesture: the grammar above requires a $\mu R$ gesture of strength less than 1, but otherwise leaves the precise value as a free parameter (see Fig. 29 in §8.4 below). Differences in the degree of C-V overlap across tokens could also contribute to token-wise variability in vowel reduction, a possibility we discuss in the next section.

**GESTURAL OVERLAP AND ‘DELETION’**. We have argued that vowel deletion in Uspanteko is in fact an extreme case of gestural overlap, corresponding to the endpoint of a gradient vowel-reduction pattern occurring in the same positions. We repeat the relevant comparison in 27.

\[ (27) \text{Vowel reduction vs. vowel deletion in } \ldots(\cdot\sigma.CV_{\mu R}C) \]

\[ (a) \text{VOWEL REDUCTION} \]

\[
\begin{array}{c}
C \quad V_{\mu R} \\
\end{array}
\]

\[ (b) \text{VOWEL ‘DELETION’} \]

\[
\begin{array}{c}
C \quad V_{\mu R} \\
\end{array}
\]

As schematized in 27, full ‘deletion’ occurs when a gesturally reduced vowel is substantially overlapped by adjacent consonants, to the extent that it is no longer audible.

Postvocalic consonants are normally coordinated antiphase with the preceding vowel, as in 27a: the articulation of the consonant begins after the articulation of the vowel has essentially completed. This differs from the timing relation in 27b, in which the postvocalic consonant begins during the articulation of the vowel itself. To produce the pattern of increased overlap in 27b, we can manipulate the COUPLING STRENGTH of

---

23 Assigning a weak $\mu$-gesture to a foot-internal syllable predicts that all consonants contained in that syllable should be reduced and/or lenited as well (Byrd & Saltzman 2003, Saltzman et al. 2008, Byrd & Krivokapić 2021). At the same time, gestural overlap between vowels and consonants sometimes seems to lead to consonant lengthening (Davidson 2006b). Hence, the predictions of our analysis for consonant duration in weak foot-internal syllables are not entirely clear. We leave this issue for future research.
C-V coordination patterns (§8.1). Coordination patterns with high coupling strength are more precisely sequenced and less variable in their timing. Conversely, coordination patterns with low coupling strength are more variable and less stable in their timing.

We assume that vowels in unstressed, foot-internal syllables are only loosely coupled with their neighboring consonants. This loose coupling is another manifestation of prosodic weakness in reduction environments. Weak foot-internal syllables should thus show more variable patterns of C-V and V-C overlap than syllables in other contexts do. This variability should especially affect postvocalic consonants, as the antiphase coordination relation specified for V-C sequences is inherently less stable than the in-phase coordination relation specified for C-V sequences (Goldstein et al. 2006, Nam et al. 2009). We implement this proposal with the two constraints in 28, and illustrate their interaction in 29.

(28) a. **Weak-Coordination[foot]**: Assign one violation for every coordination relation in a weak, foot-internal syllable with a coupling strength $CS_{Ft}$ equal to or greater than the default coupling strength $CS_D$ (*$CS_{Ft}$-weak $\geq CS_D$).

b. **Default-Coupling**: Assign one violation for every coordination relation with a coupling strength other than the default value $CS_D$.

(29) **Weak-Coordination[Foot] $\gg$ Default-Coupling**: Variably ‘deleted’ vowel in weak foot-internal syllables

\[
\begin{array}{ccc}
C & V_{\mu R} & C \\
\hline
\end{array}
\quad \sim \quad
\begin{array}{ccc}
C & V_{\mu R} & C \\
\hline
\end{array}
\]

Goldstein et al. (2006) and Nam et al. (2009) argue that in-phase coordination is the basic, default coordination pattern. For example, antiphase timing patterns can suddenly shift to in-phase timing when coordinated movements are repeated at fast rates. This is potentially relevant for our analysis: when antiphase coordination is weakly specified in V-C sequences, coordination patterns may drift toward the default in-phase timing relation, leading to even greater degrees of V-C overlap (e.g. Parrell 2012).

We again note that this analysis carries over straightforwardly from posttonic syncope in […]d.CVC to pretonic syncope in […]CV(C)ˈ. Onsets are obligatory in Uspanteko, so there are no V-V sequences (e.g. Bennett 2016). As a consequence, pretonic vowels will always be followed by a consonant—either a coda consonant or the onset of the following syllable. In either case, the same CVC coordination patterns shown in 29 will also obtain in pretonic position.

Our analysis has invoked two distinct forms of reduction in unstressed, foot-internal positions: a weak µR-gesture, and a weakening of coupling strength between gestures. Both of these mechanisms are arguably needed for an adequate analysis of Uspanteko syncope. First, simply reducing the magnitude of the vowel gesture may not guarantee enough gestural overlap to produce vowel ‘deletion’ in all circumstances. In particular, vowel ‘deletion’ routinely occurs in contexts where we expect gestural expansion, and thus reduced gestural overlap: these include utterance-final position (e.g. in isolation forms; Byrd et al. 2006, Byrd & Riggs 2008, Katsika 2012, Katsika et al. 2014) and slow, formal speech (§4.3; Davidson 2006b). Some additional mechanism is necessary to ensure that deletion can regularly occur in these contexts as well. The gestural ‘sliding’ produced by weakening coordination strength is thus a mechanism for producing extreme gestural overlap, independent of the magnitude of the vowel gesture itself.
Additionally, in §8.4 we argue that variability in gestural timing is responsible for apparent cases of ‘vowel epenthesis’ in Uspanteko that are associated with syncope (§4.9). We argued in §4.4 that the ban on deleting vowels in prefixes in Uspanteko is synchronically arbitrary and must simply be stipulated. As such, we formalize this restriction in grammatical terms, via the constraint and constraint ranking in 30.

(30) a. *µR-in-PREFIX: Assign one violation for every weak µR-gesture present associated with a prefix.
   b. *µR-in-PREFIX ≫ REDUCE-IN-Foot (defined in 24)

The ranking in 30b blocks extreme vowel reduction in prefixes, even when those prefixes occur in unstressed, foot-internal positions (the locus of syncope). This will have the effect of inhibiting syncope of prefixal vowels. For additional arguments that abstract gestural constraints may refer to morphological structure, see Bradley 2007.

8.4. Epenthesis as aggressive gestural overlap. One of the most intriguing aspects of syncope in Uspanteko is its interaction with an apparent process of vowel epenthesis (§4.9). Frequently, words with posttonic syncope are also produced with a word-final reduced vowel (31). These vowels are inserted rather than lexical: they can occur on any word with penultimate accent (31b) in which syncope occurs, and they do not occur in morphologically related forms that have final accent (31a) instead. In other words, the distribution of these word-final reduced vowels is phonologically predictable. Etymologically, these word-final reduced vowels are clearly an innovation, which further supports the claim that they are inserted rather than underlying.

(31) a. laq ['laq] ~ *[laqə] ‘plate’
   b. qálaq ['qálaq] ~ ['qálqə] ~ ['qáłqə] ‘our plate’

The quality of these inserted vowels ranges from [ə]-like to a full copy of the syncopated vowel (see again Can Pixabaj’s 2007:45–46 description of ‘metathesis’).

There are several puzzles here. First, why is epenthesis conditioned by syncope? And second, why does the quality of the epenthetic vowel vary between [ə] and a copy vowel? Our answer to both of these questions is to deny that ‘epenthesis’ as such actually occurs. Instead, we propose that syncope-dependent ‘epenthesis’ (31) instead reflects a dramatic reorganization of gestural timing in speech production.

We have already argued that syncope itself involves a high degree of gestural overlap between weak, reduced vowels and flanking consonants (Figure 27). This overlap is due, in part, to the weakening of the coupling relation between the syncopated vowel and the following consonant. Normally, postvocalic consonants are constrained to occur in an antiphase (sequential) timing relation with the preceding vowel (Fig. 27a). If this requirement is weakened or eliminated, the result will be greater (and more variable) overlap in V-C sequences (Fig. 27b). Such overlap can lead to gestural hiding, which creates the impression of vowel deletion (Fig. 27c).

But what happens as the postvocalic consonant drifts even further to the left, toward an in-phase (simultaneous) relation with the preceding vowel? In-phase timing of consonants and vowels will effectively produce a C-V sequence (§8.1): the consonant and vowel begin at the same time, but the slower, longer vowel gesture persists beyond the end of the consonant. So as a V-C sequence, which is coordinated in an antiphase pattern, drifts toward an in-phase coordination pattern, we should see a shift to C-V organization in the acoustic output. A shift from V-C organization to C-V organization is essentially the definition of metathesis—or alternatively, deletion with final copy
Epenthesis (Fig. 27; see Steriade 1990, Blevins & Garrett 1998, 2004, Yanagawa 2003, and Parrell 2012 for similar ideas).

What accounts for the variable vowel quality observed in ‘epenthetic’ vowels? We propose that the difference between [ə] and a copy vowel (i.e. metathesis) is essentially a matter of degree. If the vowel in a VC → CV reversal has an extremely reduced lingual gesture, it will manifest as [ə] (§8). If the vowel in a VC → CV reversal is less reduced, it may manifest as a copy vowel. Alternatively, if VC → CV overlap is timed such that the target phase of the vowel is obscured by the overlapping consonant, the only audible portion of the vowel may be its release phase, during which the tongue body often transitions through a neutral, [ə]-like region of the vocal tract (either to a rest position or toward a following articulation). This outcome is schematized in Fig. 27d. In either case, if the audible portion of the vowel is short enough, it may sound like [ə] simply by virtue of its short duration (e.g. Crosswhite 2001:Ch. 7).

Analyzing vowel ‘epenthesis’ as extreme gestural sliding also sheds light on apparent exceptions to the otherwise strong generalization that ‘epenthesis’ is dependent on syncope in Uspanteko. Rarely, epenthesis occurs in the absence of syncope (Figure 28, and §4.9).

![Figure 27. Vowel deletion and epenthesis as gestural sliding in kúť́’iʃ̩ ’flower’.](image-url)

![Figure 28. Gratuitous epenthesis in íŋjal ‘my corncob’ (speaker JBAT, 2011).](image-url)

Given the above gestural analysis, we can already specify the conditions under which ‘epenthesis’ would occur in the absence of syncope. Forms like those in 31 above must correspond to productions in which the vowel is timed so as to be audible both before AND after the word-final consonant. This corresponds to configurations in which a single consonant is superimposed on, and enclosed WITHIN, the full temporal extent of a longer, slower vowel gesture, as in Figure 29.\(^ {24} \)

\(^ {24} \text{The configuration we assume for ‘gratuitous’ epenthesis is very similar to the articulatory configurations that have been proposed for intrusive vowels in CC clusters in other languages; see Steriade 1990, Hall 2003, 2006, Bradley 2004, 2006, 2007, Ridouane & Fougeron 2011, and Belilik 2018.} \)
Structures like Fig. 29 can be modeled in terms of the same gestural structures and units we have already invoked in our analysis. The ‘gratuitous’ epenthesis in Fig. 28 can be produced with two adjustments to gestural planning. First, the vowel must remain relatively unreduced, despite being in a weak position. This corresponds to a µ-gesture of strength close to but still less than 1. Second, the timing relation between the vowel and following consonant must be relatively free, so that the postvocalic consonant can encroach on the vowel. This corresponds to a low coupling strength between V-C, which can lead to greater variability in timing between the vowel and following consonant.

In our OT analysis of syncope (§8), we invoked two constraints related to gestural magnitude and intergestural timing, repeated in 32.

(32) a. **REDUCE-IN-Foot:** Assign one violation for every metrical foot that does not have a µ-gesture of strength < 1 associated with its weak, unstressed branch.

b. **WEAK-COORDINATION[FOOT]:** Assign one violation for every coordination relation in a weak, foot-internal syllable with a coupling strength $CS_{Ft}$ equal to or greater than the default coupling strength $CS_{D}$ (*$CS_{Ft}$-weak $\geq CS_{D}$).

These constraints can be satisfied by a range of values for the strength of the µR gesture and V-C coupling, respectively. Let us assume that the actual values for these parameters during speaking are noisily sampled from a normal distribution, as in Figure 30, subject to the limits specified by the constraints in 32. In some cases, values for µR and V-C coupling strength will be selected such that the µR gesture only weakly reduces the size of the vowel gesture, while the coupling strength is extremely weak. These are the conditions under which ‘gratuitous’ epenthesis occurs, such as that seen in Fig. 28 above.

If the necessary parameter values for ‘gratuitous’ epenthesis occur in the tails of each distribution, as we show in Fig. 30, we correctly predict that ‘gratuitous’ epenthesis should be a relatively rare outcome.

To close, we now consider pretonic syncope. In cases of pretonic position, nothing like epenthesis or metathesis occurs. That is, we never find outcomes like *chik’oor → [tʃɪk’ɔːr] ‘hiccup’, in which the lexical vowel appears to be transposed
across a consonant. There are several possible explanations for this gap. First, a form like *[tʃiːr] would have two vowel gestures produced in direct sequence—effectively, hiatus. Hiatus is disallowed in Uspanteko, as in most Mayan languages. A constraint against hiatus, formulated in gestural terms, might therefore account for the lack of ‘epenthesis’ or ‘metathesis’ under pretonic syncope.

Alternatively, it may be relevant that forms like *[tʃiːr] would involve V-V overlap, with each vowel placing competing demands on dorsal position during the overlapping interval. If gestural ‘metathesis’ leads to V-V overlap, then the reduced vowel may be effectively obliterated because of its low resistance to coarticulation (§8.1).

In sum, we have argued that syncope in Uspanteko involves extreme gestural reduction and highly variable intergestural timing in weak, foot-internal positions. These assumptions provide a coherent explanatory account of the phonetic and phonological characteristics of syncope in this language.

9. On diagnostics for grammatical control. We began our investigation of Uspanteko syncope by applying some widely employed diagnostics for distinguishing between phonological and phonetic phenomena. In this section we reflect on which diagnostics were in the end actually probative in identifying syncope as an intentional, grammatically controlled process, rather than an incidental phonetic by-product of general conditions on speech production.

In this context, we again highlight the fact that many of the diagnostics employed in this article have been presented in the literature as tests for distinguishing phonological patterns from phonetic ones. This dichotomy fails to take into account the existence of language-specific, learned, and intentionally controlled phonetic behavior. Instead, we must ask whether any given pattern shows evidence of speaker control, and if so, ask separately whether it appears to be phonetic (i.e. gradient) or phonological (i.e. categorical) in nature.

Ultimately, several diagnostics did successfully identify vowel deletion in Uspanteko as grammatically controlled gestural overlap. The fact that deletion occurs across all speech rates and styles provides a clear indication of intentional, cognitive control. We again emphasize that the type of control involved could be either phonetic or phonological in nature, assuming that such a division is valid. For example, Gafos (2002) and Hall (2006) describe some nonphonological, intrusive vowels that persist even at slow speech rates. They analyze the lack of sensitivity to speech rate as resulting from a language-specific phonetic pattern of gestural coordination between adjacent consonants that results in an open transition at all rates of speech. We thus echo Hall’s (2006) view that sensitivity to speech rate may diagnose low-level, unintentional phonetic patterns, while lack of sensitivity to speech rate indicates some degree of speaker control, be it phonological or phonetic in character (see also Solé 2007, Dabkowski 2018, Eischens 2022). The fact that syncope occurs in utterance-final position, as well as in other environments that favor hyperarticulated, nonreduced speech, also indicates that syncope is planned and controlled.

Two other diagnostics seemed particularly effective for identifying syncope in Uspanteko as being under grammatical control. Sensitivity to metrical footing (§3, §4.10) and to morphological structure (§4.4) provides unambiguous evidence that syncope is conditioned by abstract grammatical principles and must therefore be itself a grammatically controlled process.

Somewhat surprisingly, several phonotactic conditions on syncope can be explained in terms of fairly general phonetic principles, with little or no reference to abstract
phonological constraints as such. This includes the avoidance of syncope in [ʔVC] and [CVʔ#] sequences (§4.5), antigemination effects (§4.6), and sensitivity to vowel height (§4.8). The upshot is that these phonotactic conditions on syncope provide at most limited evidence for speaker control. For related discussion, see Davidson 2006b, Kawahara & Shaw 2018, and Shaw & Kawahara 2018.

Along similar lines, the interaction of syncope with vowel ‘epenthesis’ (§4.9, §8.4) at first seemed like an indication that syncope might feed phonological processes and thus itself be phonological. However, we argued that ‘epenthesis’ as such does not actually occur in Uspanteko and should instead be analyzed as yet another consequence of gestural coordination patterns in the language. Vowel ‘epenthesis’ might thus indicate a level of speaker control, but it does not support an analysis of syncope as a categorical process of symbolic deletion.

At least one property seemed to indicate that syncope might not be a grammatically controlled process after all: syncope appears to create highly marked consonant clusters (§4.7). Retrospectively, we can see that this observation is merely a hint that syncope does not involve literal vowel deletion, but rather extreme gestural overlap. It is not actually informative as to whether gestural overlap is under grammatical control, as we have argued here.

10. Conclusion. We framed our investigation as a dilemma: is vowel deletion in Uspanteko a phonological process or a phonetic one? The answer, in a sense, is ‘both’. Syncope has the characteristics of a process that is regulated by an abstract, symbolic grammar. However, it turns out that ‘syncope’ may not involve deletion at all, at least in the strictest phonological sense. Rather, ‘deletion’ reflects language-specific patterns of articulatory coordination, implemented in physical space and time. In this way, syncope must be understood as a phonetic process, involving as it does the fine details of real-time speech production.

The resolution to our dilemma is thus synthesis: vowel deletion in Uspanteko is a grammatically controlled phonetic process. If the grammar in question is phonological in nature, then this result entails that phonological grammars must have access to abstract representations of articulatory sequencing and coordination, as in articulatory phonology and related frameworks. If the relevant notion of ‘grammar’ refers to language-specific phonetic patterning instead (e.g. Kingston & Diehl 1994), then phonetic grammars must have access to a richer array of grammatical information than is often assumed. In either case, it would appear that the boundaries between phonetics and phonology are relatively porous ones, as argued by Scobbie (2007) and others.

REFERENCES
Aylett, Matthew, and Alice Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence,
and duration in spontaneous speech. *Language and Speech* 47.31–56. DOI:
10.1177/00238309040470010201.

**Baird, Brandon.** 2011. Phonetic and phonological realizations of ‘broken glottal’ vowels in K’iche’. In Shklovsky et al., 39–50.


**Beckman, Mary E.** 1996. When is a syllable not a syllable? *Phonological structure and language processing: Cross-linguistic studies*, ed. by Takashi Otake and Anne Cutler, 95–124. Berlin: De Gruyter. DOI: 10.1515/9783110815825.95.


**Bell, Alan; Jason M. Brenier; Michelle Gregory; Cynthia Girand; and Dan Jurafsky.** 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60.92–111. DOI: 10.1016/j.jml.2008.06.003.

**Bell, Melanie J.; Sonia Ben Hedia; and Ingo Plag.** 2021. How morphological structure affects phonetic realisation in English compound nouns. *Morphology* 31.87–120. DOI: 10.1007/s11525-020-09346-6.

**Bellik, Jennifer.** 2018. *Vowel intrusion in Turkish onset clusters.* Santa Cruz: University of California, Santa Cruz dissertation. Online: https://escholarship.org/uc/item/1xn6m49b.


**Bennett, Ryan; Meg Harvey; Robert Henderson; and Tomás Alberto Méndez López.** 2022. The phonetics and phonology of Uspanteko (Mayan). *Language and Linguistics Compass* 16:e12467. DOI: 10.1111/lnc3.12467.


the mirror neuron system, ed. by Michael A. Arbib, 215–49. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511541599.008.


HENDERSON, ROBERT; TOMÁS ALBERTO MÉNDEZ LÓPEZ; RYAN BENNETT; and MEG HARVEY. 2022. Xoqoneb’: Una historia uspanteka de las tierras altas centrales de Guatemala. *TLALOCAN* 27.125–60. DOI: 10.19130/iifl.tlalocan.27.2.010087X94.


KATSIKA, Argyro; JELENA KRIVOKAPIĆ; CHRISTINE MOOSHAMMER; MARK TIEDE; and LOUIS GOLDSTEIN. 2014. The coordination of boundary tones and its interaction with prominence. Journal of Phonetics 44.62–82. DOI: 10.1016/j.wocn.2014.03.003.


Vowel deletion as grammatically controlled gestural overlap


NAM, HOSUNG; ELLIOT SALTZMAN; JELENA KRIVOKAPIC; and LOUIS GOLDSTEIN. 2008. Modeling the durational difference of stressed vs. unstressed syllables. Proceedings of the 8th Phonetic Conference of China, Beijing.


PATTERSON, DAVID; PAUL C. LOCASO; and CYNTHIA M. CONNINE. 2003. Corpora analyses of frequency of schwa deletion in conversational American English. Phonetica 60.45–69. DOI: 10.1159/000070453.

Pellegrino, François; IOANA CHITORAN; CHRISTOPHE COUPÉ; and EGIDIO MARSICO (eds.) 2009. Approaches to phonological complexity. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110223958.


PLUG, LEENDERT; Abdurrauf SHITAW; and BARRY HESELWOOD. 2019. Inter-consonantal intervals in Tripolitanian Libyan Arabic: Accounting for variable epenthesis. Laboratory Phonology 10.5. DOI: 10.5334/labphon.122.


ROON, KEVIN; PHILIP HOOLE; CHAKIR ZEROUAL; SHIHAO DU; and ADAMANTIOS I. GAFOS. 2021. Stiffness and articulatory overlap in Moroccan Arabic consonant clusters. Laboratory Phonology 12:8. DOI: 10.5334/labphon.272.


[Received 2 December 2021; revision invited 13 May 2022; revision received 22 October 2022; accepted pending revisions 7 December 2022; revision received 19 December 2022; accepted 19 December 2022]