

Chapter 6

Reconciling the debate about final obstruent voicing: The phonology of Lakota obstruent lenition

Geoffrey Schwartz^a & Shanti Ulfsbjorninn^b

^aAdam Mickiewicz University, Poznań ^bMemorial University of Newfoundland

Word-final obstruent voicing has been claimed to be unavailable as a synchronic phonological process due to universal markedness constraints. However, Evolutionary Phonology (Blevins2004) predicts the possibility of such a phonological rule, and purports to show it in various languages, touching off a debate with Kiparsky (2006, 2008). Searching for a more robust example, BlevinsEtAl2020 provide a phonetic study of voicing in Lakota. They show that Lakota has phonetic alternations between voiceless and voiced stops, with the voiced stops appearing in final position. They therefore claim the language has a synchronic phonological rule of final obstruent voicing. We call into question the assumed phonological status of final voicing in Lakota. Phonetically, Lakota does have only voiced stops in final position, but at the same time final fricatives are neutralised to the voiceless series. We show that, rather than neutralizing stops toward a marked or complex structure, Lakota has only a process of phonological lenition in weak positions: sonorisation of stops and devoicing of fricatives. Under the assumption that manner of articulation is a structural property, a single lenition process can produce final voiceless fricatives but final voiced stops.

1 Introduction

When one considers the expression ‘phonological diversity’, it is interesting to think about its sources. That is, what is it that gives richness to the phonological systems of the world? In this regard, what comes to mind first are typologically unusual or geographically restricted sound classes and/or segmental inventories. For example, a phonological use of phenomena such as non-pulmonic

airstream mechanisms or non-modal voice qualities contributes great richness to the inventory of sounds found in the world's languages. An additional source of diversity is a wide range of phonological processes across languages, ranging in scope from allophonic changes to phonologically conditioned morphological alternations. These processes show diversity both in the types of sounds that undergo changes and the types of changes that are actually observed. In this contribution, we consider what is assumed to be an unusual phonological process, namely, final-obstruent voicing, but suggest that it represents diversity not as a phonological process per se, but rather as a somewhat unusual mapping between phonology and phonetics. Unlike final obstruent devoicing, which is quite widespread (though perhaps not as ubiquitous as sometimes assumed, see Iosad, this volume), final voicing has been described in just a few languages, including Lezgian (lezg1247; Yu2004), Somali (soma1255; Blevins2006), and Lakota (lako1247; Rood2016). Here we focus on the latter of these.

Following Trubetzkoy1939, a number of modern phonological models would explain the much more common pattern of word-final or syllable-final obstruent devoicing and, conversely, the rarity of a word-final voicing rule, as the result of an Optimality Theory markedness constraint: e.g. *VOICED OBS# or *VOICED CODA (WetzelsMascaro2001; Lombardi1999). This would lead to a substantive universal, one relating to attestation rather than formal principles (Hyman2008), that can be stated as follows: word-finally one can find only unvoiced obstruents. The common assumption is that the presence of this constraint in Universal Grammar leads to an *a priori* expectation that final obstruents are voiceless, and this expectation guides language acquisition and language change (Kiparsky2006, Kiparsky2008; see also Broselow2018). The foundations of this constraint supposedly have phonetic justification – final position is associated with a decrease in the transglottal airflow required to maintain phonation (see e.g. Ohala1997).

This expectation is encoded somewhat differently in so-called representational approaches, including frameworks such as Government Phonology (KayeEtAl1990; Charette1991), Dependency Phonology (AndersonJones1974; AndersonEwen1987; Botma2004), Strict CV (Scheer2004), and Radical CV Phonology (Hulst2020). In these models, at least some traditional 'markedness' properties can be described in terms of representational complexity (Ulfsbjorninn2017), by which greater complexity is based on the number of phonological features. For example, in a true voicing language such as Polish or Russian, voiced obstruents are said to contain a feature [voice], while voiceless obstruents are claimed to be unspecified for laryngeal features (e.g. BeckmanEtAl2013). In other words, the voiced obstruents contain an extra feature, which makes them more complex. According to these models, we do not expect to find patterns of complexity reversal in positions

of neutralization. (Harris1997, Harris2005; Cyran2010; SégéralScheer2001).¹ On this view, word-final devoicing is not an instance of strengthening, but instead is a loss of a laryngeal feature, resulting in final plain voiceless stops in weak positions such as syllable codas. Under the assumption that voiced obstruents are more representationally complex than plain ones, the former should be dis-preferred in final position.

Conversely, in the approach known as Evolutionary Phonology (Blevins2004), markedness does not have any formal status. In this framework, the typological frequency of final devoicing falls out from phonetic constraints that are not formally encoded in grammar. Since there is no formal restriction on a rule which would make final obstruents voiced, and we could, therefore, expect to find such a pattern. In fact, we probably should expect to find such a pattern since, as Kiparsky2006 shows, there are multiple reasonable diachronic pathways towards this outcome. The EP perspective touched off an animated debate between Blevins and Kiparsky, since they apparently represent diametrically opposed views on the place of grammar in language acquisition and language change. In that context, the existence and phonological status of word-final obstruent voicing has become a *cause célèbre* of the debate. Blevins2004 presents a number of supposed empirical cases for word-final or syllable-final voicing. However, Kiparsky2006 argues against them, casting doubt on each one (e.g. Lezgian, Somali). BlevinsEtAl2020 (hereafter, B2020) seem to recognise this and their study on Lakota, which inspired the present contribution, is offered as a more convincing example of word-final obstruent voicing.

The putative process of final voicing in Lakota is exemplified in (??), which are taken from B2020 (p. 301). Note that in the first and third forms, prevocalic /p/ and /k/ surface as voiced [b] and [g] syllable-finally, while /t/ in the second form surfaces as [l]. The central argument of the B2020 paper (stated explicitly on p. 297) is that alternations like these in Lakota constitute a robust example of a language with a true synchronic phonological process of obstruent voicing which is supported by phonological evidence. However, as we show, the Lakota pattern apparently does not qualify as a truly phonological process. We claim

¹A reviewer calls this position into question, citing IversonSalmons2011, who summarise cases in which word-final voiceless stops may be aspirated. Thus, for example, they argue that final devoicing in German, which is said to neutralise the laryngeal contrast, is in fact a case of final fortition. While the universality of the representational perspective on final neutralization is beyond the scope of this chapter, there is phonetic evidence that Iverson and Salmons' description of German is incorrect. RoettgerEtAl2014, in a set of experiments that addressed the question of orthography, found robust evidence for incomplete neutralization of the German voicing contrast. This evidence was confirmed in a replication study (RoettgerBaer-Henney2019).

that phonologically it is a weakening process, a type of sonorisation, as reflected in the alternation between /t/ and [l] in (??), while the realization of [b, g] in (??) as fully occluded stops is a purely phonetic effect. Representational confusion between voiced stops and sonorants is expected under the assumptions of representational models in which manner of articulation is a structural property (e.g. Steriade1993, Schwartz2016), as we shall see in §4.

Table 1: Lakota word and syllable-final voicing (BlevinsEtAl2020: 301)

| | | Truncated (word-final) | Reduplicated (syllable-final) |
|-----|------------------------------------|-----------------------------------|---------------------------------------|
| /p/ | /topa/ ‘four’ | /tob/ ‘four’ | /tobtopa/ ‘by fours’ |
| /t/ | /napota/ ‘to wear out footwear’ | /napol/ ‘to wear out footwear’ | /napolpota/ ‘wearing out footwear’ |
| /k/ | /foka/ ‘to be thick’ | /fog/ ‘to be thick’ | /fogfoka/ ‘to be thick’ |

2 Lakota and final obstruent voicing

2.1 Phonetics or phonology?

If there is both acoustic and impressionistic evidence of a given phonetic phenomenon, can we assume that the phenomenon qualifies as phonological? We believe that the answer to this question is no. The velar nasal and the glottal stop in English come to mind as familiar examples. The latter is attested as an allophone of /t/ or as a boundary marker before initial vowels (see e.g. Cruttenden2001). The former is restricted in its distribution, and is sometimes analysed as an underlying nasal-stop cluster (see e.g. Smith1982). We suggest that similar forces underlie the appearance of voiced stops in Lakota. In other words, what we see phonetically is not always phonological.

The main claim that we adhere to, first advanced by Rood2016, is that voiced stops in Lakota are phonological sonorants that may be phonetically realised as stops. The phonetic voicing of stops is a result of a sonorization process. In the discussion describing Lakota and advocating the EP approach, B2020 mention Rood’s claim. However, instead of considering its merits, B2020 (p. 296) present the Lakota data with respect to a perceived dichotomy between so-called “traditional markedness” approaches and “phonetic-historical” approaches, such that it appears their goal is to undermine the former.

2.2 The status of voiced stops in the Lakota consonant system

An inventory of the Lakota consonant system, based on descriptions in B2020 (p. 299) and **Rood2016**, appears in (??). The main focus of our discussion will be on the voiced stops, since their phonological status may be assumed to be relevant to the question of whether the phonology of Lakota has a rule of final obstruent voicing. Before proceeding, however, it is worth noting that Lakota is a language with both aspiration and ejection in its stop system. In many languages with both of these features in its system of stops, voicing is not contrastive. This was apparently also the case in Proto-Siouan (**RankinEtAl1998**), which gave rise to Lakota. Similar systems, with aspiration and ejection but not voicing, are found in many other indigenous languages of the Americas (**Maddieson1984**). In South America, these include Cusco Quechua (cusc1236), Bolivian Quechua (nort2976, sout2991), Central and Chilean Aymara (cent2142, sout2996), Uru (uruu1244), Kallawaya (call1235), and Chipaya (chip1262), to name just a few (**LevEtAl2015**). In North America, these include Nuxalk (bell1234; **Nater1984**) in the Salishan family, as well as Navajo (nava1243; **McDonoughLadefoged1993**) and Apache (west2615; **GordonEtAl2001**) in the Athabaskan family, to name just a few.

Of the two voiced stops in Lakota, only the bilabial [b] has been claimed to be phonemic. The distribution of the velar [g], as pointed out by both **Rood2016** and B2020, is completely predictable. **Rood2016** compiles a list of six forms with prevocalic [b] that suggest possible phonemic status for the voiced bilabial stop in Lakota. These forms are given in (1).

- (1) Lakota words with prevocalic [b] (**Rood2016**)
 - bébelá* ‘baby’,
 - ʃkíbibila* ‘black capped chickadee’,
 - bá* ‘to blame’,
 - ábelá* ‘scattered’,
 - ka-bú* ‘to play the drum’ (ka- ‘by hitting’; bu- ‘make a hollow noise’),
 - wahíbu* ‘I left to come’.

As argued by **Rood2016**, for each of the words in (1), there is a factor that may cast doubt on whether [b] is a truly phonological entity. The first form, *bebelá*, is almost certainly a borrowing from French. The second, *ʃkíbibila*, is apparently an onomatopoetic bird name which in some dictionaries also appears with voiceless stops. The forms, *ba*, *abela*, and *kabu* apparently contain [b]-initial roots. However, the first is not widely known, and the second is shown with [p] in an earlier dictionary (**Rood2016**). The final form is described by Rood as archaic. In sum,

Table 2: Inventory of Lakota consonants, after Rood2016 and B2020.

| | labial | dental/alveolar | post-alveolar (palatal) | velar | glottal |
|---|----------------|-----------------|-------------------------|----------------|---------|
| unaspirated voiceless stops and affricates | p | t | \widehat{tj} | k | ʔ |
| ejective stops and affricates | p' | t' | \widehat{tj}' | k' | |
| aspirated stops and affricates | p ^h | t ^h | \widehat{tj}^h | k ^h | |
| voiced stops | b? | | | g? | |
| voiceless fricatives | | s | ʃ | x | h |
| voiced fricatives | | z | ʒ | y | |
| lateral sonorants | | l | | | |
| nasal sonorants | m | n | | | |
| glides | w | | j | | |

the relative rarity of the voiced bilabial stop prevocalically suggests marginal, if any, phonological status for the consonant.

Both Rood and B2020 agree that in Lakota, there is no voicing contrast in consonant clusters. An inventory of Lakota monomorphemic consonant clusters, adapted from B2020 (p. 302), is given in the table in (??). Notably, voiceless obstruents can form clusters with each other (e.g. *pte* ‘water buffalo’; *tke* ‘to be heavy’; *psa* ‘reed, straw’; examples from B2020: 302). However, in two-member clusters, when C2 is a nasal or lateral, C1 underlying voiceless stops surface as voiced stops or nasals (*ble* ‘lake’, *gma* ‘walnut’; B2020: 302), and phonetically voiceless stops cannot appear in this position. In the table below, the checkmarks indicate sequences involving no changes in consonant quality. The shaded cells indicate gaps in the cluster inventory. Finally, the clusters where voiceless stops surface as voiced are indicated with segmental symbols. What the data from cluster phonotactics clearly shows is that C1 in word-initial clusters does not contrast for voicing.

Table 3: Lakota monomorphemic clusters (C1C2), adapted from B2020 (p. 302)

| | C2 | p | t | k | tʃ | s | ʃ | m | n | l | w |
|----|----|---|---|---|----|---|---|----|----|----|----|
| C1 | | | | | | | | | | | |
| p | | | ✓ | | ✓ | ✓ | ✓ | | mn | bl | |
| t | | | | ✓ | | | | | | | |
| k | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | gm | gn | gl | gw |
| s | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| ʃ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| x | ✓ | ✓ | | | ✓ | | | ✓ | ✓ | ✓ | ✓ |

Rood2016 provides some additional cases of changes that suggest a connection between voiced stops and sonorants in Lakota. First of all, there is a rule by which stop+[l] clusters become nasalised when they precede a nasal vowel: *bla* + *ĩ* > *mnĩ* (*kte*) ‘I will go’, cf. *bla* ‘first person agent’ (**Rood2016**). Similarly, stops may become allophonically nasalised if they are preceded by a nasal vowel, and are variably realised as: [b, ^mb], [^mb] or [m], [g], [^ŋg], [^ŋg] or [ŋ] (**Rood2016**). Finally, **Rankin2001**, cited in **Rood2016** observed that the glide /w/ is found before any vowel of the language, except /u/. However, /wu/ sequences instead surface as [bu], although, as reported by **Rood2016**, no examples are given. This pattern

may be seen to support the claim that the [b] is an underlying sonorant. In attempting to explain the distribution of /w/, positing underlying /b/ would require a rather unnatural rule weakening /b/ to [w] before all vowels except /u/.

While the discussion above raises questions about the phonological status of voiced stops in Lakota, it is worth noting that beyond Lakota, sonorisation of stops is quite a common phonological process. The results of sonorisation can be quite phonetically diverse, ranging from full vocalisation: /p, t, c, k/ [β, ɶ, j, ɥ] in Maxakalí (maxa1247; see Gudschinsky, **GudschinskyPopovich1970**, **Silva2015**; **SilvaEtAl2020**), to consonantal sonorants, as well as fricatives. In fact, in many systems one gets mixed outcomes with both sonorants and fricatives resulting from weakening: /p, t, k, d/ > [ɸ, s, x, ʁ] (Ontena Gadsup (gads1258); **FrantzFrantz1966**; **Frantz1994**), or: /b, p, t, q/ > [w, f, r, x] (Seereer Siin, **McLaughlin2000**).² Each of the cases discussed here may be suggested to fall into the category of lenition, or weakening. Stop weakening is often described as ‘spirantisation’ and it produces fricatives. Often, however, these purported fricatives have such weak noise that they might be better described as approximants. Sonorisation, therefore, as a product of stop lenition is probably even more common than is reported. Spanish ‘spirantisation’ is an example (**Salinas2015**), since the ‘fricatives’ [β ɸ ɣ] it produces lack robust aperiodic noise and are difficult to distinguish from approximants.

In sum, these cases constitute evidence in support of a strong representational connection between stops and sonorants, which we suggest is exemplified in Lakota. A phonological sketch of that connection is presented in §4. At the same time, it is worth noting that sonorisation of stops is incontrovertibly attested in Lakota, since [l] surfaces as the outcome of the ‘voicing’ of final /t/ (see the 2nd example in (??)). Therefore, our basic claim is that the phonological change that B2020 refer to as ‘voicing’ of final /p/ and /k/ is better characterised as sonorisation, even if it results in phonetic stops. In other words, we will build on Rood’s position that [b], [l], and [g] in Lakota are phonologically sonorants.

At first glance, this appears to be the kind of phonology/phonetics mismatch that is common and expected in Substance Free approaches to phonology (**HaleReiss2008**; **Blaho2008**; **ChabotScheer2019**). These approaches argue against the incorporation of phonetic considerations into models of phonological representation or computation. Since [b l g] apparently constitute a phonetically unnatural class, their parallel behaviour in Lakota might be taken as evidence in support of such an approach.

²The most distinct aspect of obstruents is their potential for voice contrasts. Laryngeal contrast in sonorants, as in Burmese (**BhaskararaoLadefoged1991**), is quite uncommon (**Maddieson1984**).

At the same time, in §4, we show that from the perspective of models in which segments are split into smaller structural entities (e.g. Steriade1993; Schwartz2016), the phonological commonalities of Lakota [b l g] may be linked to explicit phonetic properties. Phonologically, all three of the sounds may be unified as redundantly voiced segments with occlusion, whereas the noise bursts of [b] and [g] are phonetic details that are not encoded in the representation. In the case of [b], the exclusion of the release burst may stem from a phonetic universal by which labially produced noise is weak in amplitude, hence the typological rarity of bilabial fricatives (LadefogedMaddieson1996). In the case of [g], the exclusion of the burst may stem the fact that the noise spectrum of dorsal releases is highly context-dependent (Stevens1998). For the phonology of Lakota, it may therefore be preferable to treat the burst as a redundant feature, rather than attempt to interpret its acoustic variability.

2.3 Final obstruents in Lakota – a unified process of weakening

At this point we proceed to one of the more interesting aspects of final obstruents in Lakota – the asymmetrical behaviour between final stops and final fricatives – an issue that B2020 do not attempt to explain. Fricative obstruents are unambiguously contrastive for laryngeal features prevocally, as shown in (2). However, this contrast is neutralised in coda and final positions in truncated forms in (3) (some with reduplicated roots), where fricative obstruents surface phonetically as voiceless (examples from B2020, p. 300). A comparison of the examples in (3) with those in (??) makes it clear that while final stops in Lakota apparently show voicing, final fricatives undergo devoicing.

- (2) /si/ ‘foot’ /zi/ ‘yellow’
 /ʃota/ ‘smoke’ /leʒa/ ‘urinate’
 /xolya/ ‘being gray’ /yopa/ ‘snore’
- (3) /tʃhaya/ ‘ice’ /tʃax/ ‘ice’
 /leʒa/ ‘urinate’ /leʃleʒa/ ‘urinate often’
 /koza/ ‘wave’ /koskos/ ‘wave’

Since stop voicing and fricative devoicing happen in the same position, it is desirable to provide a unified phonological explanation for both cases.

The first step in this explanation is to adopt an assumption that both stop voicing and fricative devoicing occur in what might be considered a phonologically weak position – the coda of a syllable, either word-finally, or in the C1 position

of consonant clusters. The unity of these positions may be envisioned in a Strict CV approach (SégéralScheer2001; Scheer2004), which posits that phonological structure is made of an alternating string of onsets and nuclei (C and V slots). Under this view, all consonants are obligatorily followed by a V-slot, which ‘licenses’ them. However, in the case of non-prevocalic consonants, V-slots are empty and cannot ‘license’ their onsets, which renders them weak. In this way, the commonalities between syllable codas in word-final position and C1 in two-member consonant clusters fall out directly from the postulates of the model. In each case, an onset is ‘unlicensed’. More traditional approaches cannot provide this unity – it must be stipulated, since they contain no representational link between clusters and codas.

A Strict CV visualization of phonological weakness is shown in (??). There we see a unification of environments where the obstruent’s laryngeal feature ([fortis]; for discussion of this choice of specification, see §3) is permitted when licensed by an adjacent filled V-slot. As we see in (??), [fortis] cannot be licensed before an empty V-slot and it is consequently deleted, which is shown by the crossed-out association lines.

Figure 1: Laryngeal specification licensing – a visualization in the Strict CV framework (T indicates ‘truncated’, as shown in (??) and (3))

In Strict CV, when a phonological feature is licensed, it may spread to unlicensed positions. This is shown in the representation of two-member clusters in (??). C1 is unlicensed since it is followed by an empty V-slot, and thus cannot support a laryngeal contrast. Meanwhile, C2 is licensed by the filled V-slot to its right. Consequently, the C2 [fortis] specification of /t/ in *pte* ‘water buffalo’ is in a strong position and is allowed to spread to /p/. By contrast, in *ble* ‘lake’ the C2, a sonorant, has no laryngeal specification, while the C1 stop is unlicensed, so the entire cluster is cannot support a laryngeal feature.

In looking at (??), the reader might notice that removing [fortis] results in voicing in stops, but devoicing in fricatives. We take up the question of how this is possible in §3 and §4. In §3, we show how to characterise the laryngeal specification properly, while in §4 we show how weakening leads to devoicing in fricatives and voicing in stops. In (??) and (??), we merely offer a visualization of phonological strength, for which Strict CV is well suited. At the same time, Strict CV is not well-equipped to deal with the asymmetry between stops and fricatives. For this purpose, in §4, we adopt the representations of the Onset Prominence framework (e.g. Schwartz2016), in which manner of articulation is represented structurally (cf. Steriade1993).³

³Reviewers and editors of this chapter have commented on the fact that we use two separate rep-

Figure 2: C1C2 clusters and lenition

In the meantime, it is worth noting that if we accept B2020’s interpretation and posit final voicing as a phonological rule, we would need two separate and contradictory rules to handle the processes, one of fricative devoicing and one of stop voicing. By contrast, in (??) and (??), we show how these processes may be unified as loss of a specification in weak position. The table in (??) summarises our perspective on phonological strength and weakness in Lakota. Voiceless stops weaken to what might be called “hard” sonorants (pronounced as stops in the case of the labial and velar), while fricatives lose their laryngeal contrast in favour of the voiceless series.

Table 4: Lakota consonantal lenition – strong and weak positions

| Strong | | Weak | |
|----------------------------------|-------------|------------------------------------|-------|
| Laryngeal contrasting fricatives | s ʃ x z ʒ ɣ | Neutralised fricatives - voiceless | s ʃ x |
| Voiceless stops | p t k | “Hard” sonorants | b l g |

3 Voiced can be unmarked

One of the most provocative and insightful aspects of Evolutionary Phonology is the elimination of markedness as an element in phonological grammars. The term ‘markedness’, of course, has been used to mean many different things (for discussion, see Hume2011). The meaning excised from Blevins’ model is that of ‘less optimal’. That is, for Evolutionary Phonology, phonological processes are not goal-oriented — they do not serve to make sound systems more optimal, contrary to what is assumed in theories such as Optimality Theory (PrinceSmolensky1993).

representational models (Strict CV and OP) to explain the Lakota voicing data. There are number of arguments why this choice has been made. First of all, Strict CV would require stipulation to deal with the manner-based asymmetry, a phenomenon for which OP is well-suited to describe. At the same time, OP’s phonotactic mechanisms would be required to explain the cluster data, and these mechanisms are sufficiently complex so as to require a much longer introduction than would be desirable for a volume such as this. Strict CV is more economical for showing a unified view of phonological weakness. At the same time, it should be mentioned that Strict CV and OP share the important assumption that all consonants are “onsets” by default, and that this onset status is built into their representations, rather than being derived by some sort of “syllabification” rules.

Without markedness, Evolutionary Phonology suggests that there is no *a priori* expectation for final voicing to be any ‘worse’ than final devoicing. Rather, the typological rarity of final voicing is attributed to phonetic factors that work against the evolution of such a pattern. For the most part, we are sympathetic to this outlook – markedness is a concept that we feel has been overused (again, see **Hume2011**), and phonetics may indeed explain a lot of sound patterns. For that reason, in the privative frameworks in which we work, markedness *per se* has no formal status. Instead, a more explicit designation is employed – a given entity may be phonologically specified or unspecified. This strategy makes no reference to typological frequency, difficulty in production, or other factors that are traditionally associated with markedness, but are often contradictory in nature (again, see **Hume2011**). We therefore use the term “unmarked” synonymously with “unspecified”, raising the question of what an unspecified phonological object may look like phonetically. In the case of obstruent voicing, a mainstream tradition (see e.g. **BeckmanEtAl2013**) tells us that voiceless is unspecified. Here we offer arguments that the opposite may be true.

Assuming that phonological specifications reflect observable phonetic properties, a natural conclusion is that voicing during the production of obstruents should reflect the presence of a phonological feature [voice]. In this sense, any systemic change in whether phonation is observed, be it voicing or devoicing, should constitute a phonological process. If we are to accept this proposition, final voicing in Lakota would certainly qualify for phonological status. However, when talking about ‘observable phonetic properties’, we need some kind of a reference point. If our reference point is silence, then periodicity in stop closures is clearly an observable property. A question which is not often asked, however, is whether silence is an appropriate reference point. Silence is a property of voiceless stop closures, which act as crucial acoustic ‘landmarks’ for listeners in parsing the signal to access linguistic content (see **Stevens2002**). Therefore, it is reasonable to suggest that silence is an observable phonetic property itself, rather than a reference point.

An alternative view of acoustic reference points for phonological features is offered by Traunmüller’s **Traunmüller1994** Modulation Theory, in which linguistic content, including phonological features, is encoded as modulations of a carrier signal. According to Modulation Theory, the ideal carrier is a voiced, schwa-like vocoid with evenly spaced formants. This vocoid serves as acoustic background for the transmission of an utterance and bears phonological features, as well as extra-phonological information such as the speaker’s age, sex, or emotional state. Since the carrier is voiced, periodicity by itself, at least during stop closure, is not a modulation, and voicing is not a good candidate for phonological

status. If this idea is taken to its logical conclusion, we must assume that voicing processes in stops entail the loss of specification.

The claim that voiced consonants are phonologically unspecified is widely accepted for so-called ‘aspirating’ languages such as English and German (IversonSalmons1995). In these languages, it is aspirated stops that are assumed to bear a phonological feature. By contrast, for so-called ‘voicing’ languages such as Polish or French, many scholars (see discussion in e.g. BeckmanEtAl2013) assume that voiceless obstruents are unmarked, and that voiced obstruents are specified with a feature [voice]. At the same time, however, this question is far from being settled. In particular, a number of recent proposals have argued that fully voiced obstruents are unspecified in Hungarian (Blaho2008), Breton (Iosad2012), Southwest Polish (Cyran2014), Dutch (Hulst2015), and Italian (BalognéBérceHuszthy2018).

The representational claim in the proposals mentioned above is compatible with phonetic evidence on the perception of voice contrasts in voicing languages. For example, in languages such as Polish and Dutch, a lack of phonetic voicing typically does not induce voiceless percepts (AlphenSmits2004; SchwartzArndt2018; SchwartzEtAl2019). In this connection, VOT (voice onset time) category boundaries between /p t k/ and /b d g/ are often safely above zero. In Dutch, Spanish, and Japanese (FlegeEefting1987a, FlegeEefting1987b; WilsonHashimoto2013), for example, a voiceless stop with VOT of around 20 ms, but with F0 and F1 cues matching those of voiced stops, is not in fact perceived as voiceless. If the contrast were based on [voice] and VOT, we would expect the category boundaries of zero, and all items without pre-voicing to be perceived as voiceless. Additionally, in two-series laryngeal systems, equivalence classification (Flege1987), by which bilinguals and L2 learners confuse phonetic categories between languages, is consistently greater for the voiced series than for the voiceless series (see Schwartz2022 for discussion). Taken together, these findings suggest that lenis stops are phonologically equivalent in voicing and aspiration languages, regardless of whether they are in fact voiced. Since no one denies that lenis stops in aspirating languages are unspecified, we must conclude the same for fully voiced stops.

A question that immediately arises with regard to this claim concerns cases of regressive voicing assimilation, which is attested only in languages with fully voiced stops (van RooyWissing2001; WetzelsMascaro2001). Textbook treatments (see e.g. GussenhovenJacobs2017) of this phenomenon assume a process by which a [voice] specification spreads from C2 to C1 in the cluster, implying that [voice] must be an active phonological specification. In other words, the question may be formulated as follows: how can you have regressive voicing without a feature [voice]?

There are a number of problems with the textbook interpretation that can be gleaned from a careful phonetic study of the process. First of all, based on findings from Dutch and Hungarian (Jansen2004) it appears that the neutralization that is said to occur from regressive voicing is phonetically incomplete. Even when phonetic voicing is present in these contexts, the underlying voicelessness of C1 is observable in other acoustic parameters such as F0 and burst amplitude. If regressive assimilation took place according to the textbook depictions (such as GussenhovenJacobs2017), we would expect complete neutralization in these contexts, wiping out all the cues to voicelessness, rather than just the lack of voicing in the closure. An additional fact about regressive voicing contexts is that the actual source of voicing in C1 of obstruent clusters is the preceding vowel, rather than C2. That is, voicing always ‘bleeds’ (see Davidson2016), or spills over, from the preceding vowel into the cluster. An illustration of voicing bleed is shown in an acoustic display in Figure 4 below. The figure shows a case of putative regressive voicing assimilation (t[F0E0?]/d/_d) in Polish *zbyt długo* [zbɪd dwugɔ] ‘too long’ (after Schwartz2019). In the figure, it is clear that there is phonetically more voicing in C1, the supposed ‘target’ of assimilation, than there is in C2, the supposed ‘trigger’. The closure of C1 is fully voiced, while voicing bleed from the preceding vowel ceases about halfway through the closure of C2. Phonetic studies that have compared the amount of voicing in C1 vs. C2 in assimilation contexts (HalléAdda-Decker2011; Schwartz2019) have consistently found more voicing in C1 than in C2.

Figure 3: Acoustic illustration of regressive voicing in Polish *zbyt długo* [zbɪd dwugɔ] ‘too long’ (after Schwartz2019). The top tier shows closures of final /t/ and initial /d/. The bottom tier shows the voiced portions of those closures.

The aforementioned textbook representation of regressive voicing implies a claim that [voice] spreads from right to left. However, what is clear from the studies described above and visible in the acoustic display in Figure 4, is that in putative regressive voicing contexts, phonetic voicing spreads from left to right.⁴ Thus, if we accept the textbook account, we are forced to claim that phonetic and phonological voicing processes show completely opposite directionality. While we accept that phonology can and does undo phonetic processes, we are not

⁴This is to be expected considering the aerodynamics involved in phonation — the constrictions of two consonants lead to an increase in pressure in the supralaryngeal cavity, hindering transglottal airflow (see e.g. Stevens1998, or any textbook that discusses the aerodynamics of phonation). As a result, as pressure increases, phonation ceases gradually and cannot restart until the constriction is released.

comfortable with the claim that phonology can reverse the direction of a phonetic process. The textbook representation of regressive [voice] spreading, however, implies that we accept just such a claim. In sum, the phonetic evidence, and the desire to avoid controversial claims about the phonetics-phonology relationship, suggest an interpretation that regressive voicing is a phonetic rather than a phonological process, despite the conventional accounts of phonological spreading.

The discussion above, which may at first glance seem like something of a digression, is intended to provide insight into phonetic voicing processes more generally. In particular, we have shown how basic phonetic principles can support a claim that phonetically voiced stops need not be phonologically voiced. Regardless of the status of [voice] in voicing languages, Lakota appears to have more in common with aspiration languages, as pointed out by **Rood2016**, since aspiration is contrastive in its stop system, while voicing for the most part is not. There is, therefore, no reason to assume that [voice] is active in the phonology of the language. Rather, it appears as if obstruents are specified for a feature denoting voicelessness. Lakota is unusual in that this feature produces phonemic contrast in fricatives, but apparently not (or only marginally) in stops. Without a feature [voice], the phenomenon of final voicing, described by **BlevinsEtAl2020**, cannot qualify as a phonological process of voicing. It must be something else. To complete the story about Lakota, we need to examine the representational connections between obstruents and sonorants, which is the focus of the following section.

4 The phonological structure of consonants in Lakota

Since we have suggested that a feature [voice] may be absent from two-series laryngeal systems, despite the well-known opposition between voicing and aspiration languages, an obvious question that arises is how to account for the voicing-aspiration distinction without recourse to a feature [voice]. A possible solution is offered by representational models in which obstruents are divided into smaller structural units, such as Aperture Theory (AT; **Steriade1993**) or the Onset Prominence framework (OP; see **Schwartz2016**, **Schwartz2017**). In both of those models, a stop consists of multiple positions corresponding to the closure and release phases in their production. One of these positions, derived from the transition from a stop to the following vowel (A_{\max} in AT; Vocalic Onset, or VO, in OP), is voiced by default. In essence, this position may be interpreted as part of the acoustic carrier signal, as envisioned in Modulation Theory and discussed in the previous section. Voicing during closure in stops, when it occurs,

need not denote the presence of a phonological feature, but rather the emergence of the carrier in the absence of a salient modulation, i.e. a feature denoting voicelessness. Meanwhile, unvoiced /b d g/ in aspiration languages, such as English, are attributable to aerodynamic constraints on phonation. In most aspiration languages (but see HelgasonRingen2008 on Swedish, or JacewiczEtAl2009 on Southern US English), there is no motivation to overcome these constraints, since the timing of the laryngeal feature induces long VOT and a perceptually robust contrast even in the absence of voicing (see Schwartz2017, Schwartz2022 for further discussion of two-series laryngeal typology).

Now we are ready to describe voicing phenomena in Lakota. This will be done using Onset Prominence representations. Before proceeding, we offer a brief introduction to the model; see Schwartz (2016, 2017) for a more thorough presentation. Onset Prominence representations are derived from a hierarchy of phonetic events associated with a stop-vowel sequence, typologically the most common ‘syllable’ type across languages. The fundamental building block in OP is, therefore, a prosodic unit, a stop-vowel CV that provides the material from which ‘segmental’ representations may be constructed. The stop-vowel hierarchy is shown in the tree structure on the left in (?). Each layer of the tree on the left in (?) is labeled for the phonetic event in the stop-vowel sequence from which it is derived. At the top of the hierarchy is Closure (Closure; C), the defining property of stops. The next level down is Noise (N), which is derived from aperiodic noise associated with stop release bursts, affrication, aspiration, and frication. Below Noise is the Vocalic Onset (VO) level, derived from the CV transition in the stop-vowel sequence and reflecting its periodicity and formant movement. At the bottom of the hierarchy is the Vocalic Target (VT) node, which encodes the more or less steady formant targets associated with vowels.

Figure 4: The Onset Prominence hierarchy (left) and OP manner categories (the following five graphs)

The structures in (?) reveal the OP perspective on the relationship between prosodic and segmental units in phonology. From the tree on the left, categories of manner of articulation are derived, as shown in the remaining five trees. These categories are defined in terms of binary nodes which encode the phonetic events from the stop-vowel hierarchy that are present in the articulation of a given segment type. By contrast, unary nodes appear when the event associated with the given level of the hierarchy is absent. For example, stops contain closure, noise, and release, so their top three layers are binary in (?). Nasals lack noise bursts, so their Noise node is unary, fricatives lack closure as shown by their unary C node, and so on.

In the OP framework, place and laryngeal specifications are encoded as feature annotations to the terminal nodes of the tree structures. As mentioned in the previous section (see also Schwartz2017), the laryngeal specification we assume is a single [fortis] feature. Therefore, in a two-series ‘voicing’ language, the opposition between /b d g/ and /p t k/ would be that the latter contain [fortis], while the former are unspecified. As a result, the annotation would appear as [fortis]/Ø on the VO node.⁵

Now we may turn to the link between stops and sonorants in Lakota. This is shown in the representations in (??). They first show the structure of stops in most languages, where stops contain Closure (C), Noise (N), and Vocalic Onset (VO) nodes, second, sonorant-like stops in Lakota, and third, nasals/laterals as they appear in most languages.⁶

Figure 5: OP structures for stops in typical voicing languages, Lakota stops, and nasals/laterals

For OP, the difference between stops and sonorants lies in the status of the Noise node. In sonorants, the node is absent, and in the representations in (??), it is absent from phonetic stops in Lakota. Therefore, the question that must be addressed is how an integral part of the phonetic production of stops, the noise burst, may be eliminated from the representation. A simple phonetic explanation presents itself. Stop release bursts are transient acoustic events that are extremely short in duration (Stevens1998). This is in contrast to other aperiodic noise events following stop release, including frication, aspiration, and affrication, which last much longer. Thanks to this transience, we can identify a clear phonetic underpinning to the claim that non-aspirated stops in Lakota may lack a Noise node — the bursts may not be acoustically robust enough to be represented.⁷ Additional motivation for this postulate may be found in light of the

⁵Although the representation of Lakota aspirates and ejectives is orthogonal to the discussion in this paper, we may speculate about how best to represent them. Aspirated stops, would assign [fortis] to the Noise node, while ejective stops would either assign [fortis] to the Closure node, or require an additional [constricted glottis] feature on the VO node.

⁶In OP, whether or not laterals contain the Closure node corresponds to the variable status of /l/ with respect to the feature [continuant]. This point is orthogonal to the discussion here, since in OP it is the Noise node that is responsible for the sonorant-obstruent distinction.

⁷This postulate, in turn, raises an additional question of why stops in most languages are assumed to contain the Noise node (cf. Aperture Theory, in which stops do not contain A_f corresponding to frication.). One answer to this question is that the default phonetics-phonology mapping in OP is transparent — stops without the Noise node present an opaque mapping and should be uncommon. Another answer concerns the representation of spirantization processes by which stops are lenited into fricatives — this process may be represented in OP as the loss

fact that Lakota also contains aspirated stops and ejectives, which produce more robust aperiodic noise. The prominent noise associated with aspirates and ejectives may serve to reduce the perceptual robustness of the noise associated with non-aspirated pulmonic stops.

The postulate that stop release bursts may be ignored for phonological representation is not as radical as it may seem. For an analogical case, consider voicing in vowels. In most languages, specific aspects of voice quality (stiffness, slackness, creak, etc.) are not perceived as part of a vowel's phonological representation in terms of location on a two-dimensional F1/F2-based vowel chart. Yet vowels across languages are overwhelmingly voiced. Thus, we may assume that voicing itself is a phonetic enhancement mechanism, which renders the formants that define vowel quality more perceptually robust. Likewise, we suggest the need for phonetic enhancement may turn Lakota sonorants into stops.

With the stop-sonorant connection in place, we consider what happens in final position in Lakota, which is also somewhat unusual from the OP perspective. The VO node in OP representations derives from the CV transition of a stop-vowel sequence. In final position, we may expect this node to be absent. This gives rise to neutralization, since VO is the node which houses the feature denoting a laryngeal contrast. Alternatively, if the laryngeal contrast is maintained, we may assume the VO node has been reconstructed (see SchwartzEtAl2021). Voicing of stops in Lakota arises from a slightly different process. Instead of neutralization targeting the structural node, it removes the [fortis] feature, but leaves the VO node intact. What is left is representationally equivalent to the sonorant in (??), and is phonetically voiced. This is shown in representations in (??).

Figure 6: Initial vs. final position for stops and nasals/laterals: Lakota vs. most languages

OP's structural interpretation of manner of articulation provides insight into Rood's claim that voiced stops in Lakota are in fact sonorants, which is reflected in the fact that final /t/ is voiced to [l] and not to [d]. Importantly, the OP perspective shows how confusion can arise between stop and sonorant categories, which in OP are constructed from the same representational materials. The phonological link between stops and sonorants is evident in the representations in (??) and (??). Each contains Closure and VO nodes, so it is a simple matter to envision

of the Closure node (e.g. Schwartz2016: 45), while the fricative that remains of course contains Noise. It is for this reason that OP is preferable to Aperture Theory (Steriade1993), in which A_f is absent from the representation of stops. That is, in AT, stops contain no aperiodic element, which greatly complicates the representation of spirantization processes, in which both input and output contain such an element.

confusion between voiced stops and sonorants. Indeed, as discussed earlier, the sonorization of stops is a commonly attested process.

At the same time, assuming we are correct and (unaspirated and pulmonic) stops in Lakota are indeed phonologically equivalent to sonorants, it is worth considering possible explanations or interpretations for what appears to be a relatively unusual mismatch between the phonological representation and the acoustic signal. One possibility, mentioned earlier, is that the release bursts associated with the stop productions constitute a phonetic enhancement device aimed at preserving the structure of the consonant. In final voicing contexts, this may be thought of as a compensation for the loss of the [fortis] specification. In other words, producing non-phonological release bursts prevents further weakening, or even deletion, of the consonant. In this connection, it is worth emphasizing once more the fact that /t/ is voiced to [l] rather than to [d]. Weakened coronal stops have a tendency to be realised as flaps, non-sibilant fricatives, or weak approximants, which from the perspective of OP would presumably entail the loss of the Closure (and sometimes Noise) node. The lateral realization serves to preserve the (medial) Closure of a coronal stop, which would otherwise be susceptible to weakening. This step is of course not feasible in the case of labials and velars, so they surface as voiced stops to preserve their structures shown in (??) and (??).

In the representations in (??), we also gain some insight into the OP perspective on phonological strength and weakness, which is somewhat more intricate than the ‘licensing’ approach in Strict CV described in §3. In essence, the larger the structure and the more feature annotations it contains, the stronger it is. A prevocalic stop (the leftmost tree in (??)) encodes the CV transition as the VO node. This node may be absent from final and pre-consonantal positions (the second tree from the left in (??)). The loss of the VO node may be thought of as a form of weakening. An alternative option in the model posited in (??) for Lakota (the third tree from the left), is that a feature specification may be lost but the structural node is preserved. In (??), therefore, we can see that there are multiple mechanisms that may be described as weakening in OP, each with a different output.⁸

The final aspect of Lakota voicing neutralization that must be accounted for is the asymmetry between stops and fricatives. This is shown in the representa-

⁸A third mechanism, not shown in (??), is to trim the tree from the top, turning binary nodes into unary ones. For example, taking a binary Closure node and making it unary provides a natural expression of spirantization processes that are common in intervocalic position. Since we are dealing with final position here, a thorough discussion of intervocalic position is beyond the scope of this paper.

tions in (??). Note that stops and fricatives in Lakota in (??) differ with regard to whether the VO node is present. Recall that VO in OP representations derives from the CV transitions of consonants in ‘onset’ position. These transitions are crucial for consonant perception in the case of stops, considering the transient acoustic nature of stop releases (Wright2004). By contrast, fricatives contain robust ‘internal’ acoustic cues to their identity (Wright2004). As a result, listeners are less reliant on the CV transition for fricative identification, and so their phonological representation is more likely to lack the VO node. Because of this, both initial and final fricatives in Lakota contain only a Noise node, which can and does support a laryngeal contrast.

Figure 7: Stops vs. fricatives in Lakota

The representational postulates in (??) explain the asymmetry between stop voicing and fricative devoicing in Lakota. Unlike VO, which is voiced by default since it derives from a voiced part of the signal, the Noise node is by default voiceless. The reason is that a constriction sufficient to induce frication noise hinders airflow through the glottis and voicing. In this connection, if we consider Modulation Theory’s notion of an acoustic carrier from the perspective of OP, we can arrive at a situation in which laryngeally unmarked stops can be voiced since they contain VO, but unmarked fricatives can be voiceless because they do not. As a result, the loss of the [fortis] feature surfaces as voicing in stops with a preserved VO node, but devoicing in fricatives. To make such an analysis possible, we must assume that the [fortis] feature is an abstract entity, whose phonetic realization is dependent on the level of the OP hierarchy at which it occurs.

The take-home message from the phonological perspective presented in this chapter is that the neutralization patterns observed in Lakota may be thought of in terms of weakening or loss (cf. Harris2009) of a representational primitive. We are not talking about phonological voicing or devoicing processes here, since we do not posit a [voice] feature that corresponds to any unified phonetic correlate. Rather, the phonetic effects of neutralization differ as a function of a manner of articulation, which is represented structurally in the OP model.

A final point to consider in this discussion is how more traditional approaches might handle the Lakota facts. If we assume, as B2020 do, that final voicing in Lakota is a process that is phonological in the traditional sense, an Optimality Theory analysis would have to posit an empirically unmotivated constraint *VOICELESSCODA, which outranks *VOICEDCODA, but only when the obstruent is a stop and not a fricative. Clearly, there is no way such an analysis could work

without a great deal of questionable stipulation. Rood's solution to this problem, as outlined above, is based on a classification of Lakota stops as sonorants, a claim with which we agree. Our primary critique of Rood's solution is that he provides no explanation of the origins of the representations he adopts. That is, while observing the phonological connection between stops and sonorants in Lakota, he offers no story about how stops and sonorants may be linked representationally.

5 Conclusion

This chapter has shown that, in Lakota, both final voicing of stops and final devoicing of fricatives may be easily analysed as the result of a weakening process in which a phonological specification is lost. Our fundamental claim in this regard is that the feature [voice] is not active in the phonology of Lakota. Therefore, what B2020 observe in their phonetic study cannot be not a case of phonological voicing. A more fundamental contribution of this chapter is the perspective it provides on how mismatches can arise between phonological category membership and phonetic realization. Such cases often involve attributing phonological status to something that is absent in the acoustic signal, a strategy which underlies e.g. the commonly-invoked 'empty nucleus' postulate. Final voicing in Lakota, however, requires the opposite strategy — suggesting that something present in the acoustic signal is not a phonological entity. Pulmonic stop release bursts, which are present in the acoustic signal, appear to have no phonological status in Lakota.

It has been also shown in this chapter, we hope successfully, that when a given phonetic property lacks phonological status, it need not be the case that there is a mismatch between phonetics and phonology. In other words, phonetic considerations may in fact underlie the non-phonological status of a given phonetic property. We suggest that pulmonic stop release bursts in Lakota, instead of comprising part of the phonological representation of a consonant, constitute a phonetic enhancement device to preserve a weakened consonant from further weakening or deletion. This hypothesis underlies the stop-sonorant connection that explains both the final voicing of Lakota stops and the difference in behaviour between final stops and fricatives in Lakota. More generally, we suggest that a deeper understanding of phonological diversity in the languages of the world requires an approach based on refined phonological representations that directly encode positional and contextual effects, as well as perceptual ambiguities found in the acoustic signal.

List of abbreviations

| | | | |
|-------|--|-----|---|
| AT | Aperture Theory | OP | Onset Prominence |
| B2020 | BlevinsEtAl2020 | VO | Vocalic Onset node (in Onset Prominence framework) |
| C | Closure node (in Onset Prominence framework) | VOT | voice onset time |
| N | Noise node (in Onset Prominence framework) | VT | Vocalic Target node (in Onset Prominence framework) |