*J. Linguistics* **59** (2023), 531–575. 
doi:10.1017/S0022226722000421

# Onset conspiracy in Upper Sorbian[1]

JERZY RUBACH

*University of Iowa and University of Warsaw*

(Received 5 March 2020; revised 8 September 2022)

This article has two goals: descriptive and theoretical. On the descriptive side, the article presents a grammar of gliding and epenthesis of Upper Sorbian. The descriptive goal is worthy because Upper Sorbian has a highly complex but regular and productive system of gliding and epenthesis. Upper Sorbian stands out from a typological point of view because it has ten [*sic*] different strategies to satisfy ONSET.

On the theoretical side, the question is whether Optimality Theory that has been designed to solve conspiracies can deal with the complexities of Upper Sorbian. The answer is that it cannot unless it is modified to admit derivational levels. A point of interest is that level 1 in Upper Sorbian must be defined as the root level, not as the expected stem level that includes roots and affixes. Further, it is demonstrated that Itô and Mester's CRISP EDGE constraint makes wrong predictions for Upper Sorbian, so a new constraint, MULTI, is postulated. Also, the analysis bears on the issue of positional markedness VERSUS positional faithfulness and the question of whether Duke of York derivations should be admitted in phonology.

KEY WORDS: cycle, Derivational OT, derivations, Duke of York, Optimality Theory, phonology, Stratal OT, Upper Sorbian

This article has two goals: descriptive and theoretical. On the descriptive side, the article presents a grammar of gliding and epenthesis of Upper Sorbian, a language whose literature in the generative framework virtually does not exist.[2] The descriptive goal is worthy because Upper Sorbian is a *Paradebeispiel* of a complex but regular and productive system of gliding and epenthesis. Upper Sorbian stands out from a typological point of view because it has ten [*sic*] different strategies to satisfy ONSET, a constraint that enforces the derivation of CV syllables. What we see is a

[1] I would like to thank the anonymous *Journal of Linguistics* reviewers and the associate editor for their discussion and criticism which led to considerable improvement of both the content and the presentation of my analysis. My special gratitude goes to my Upper Sorbian native speakers and consultants (in alphabetical order): Helena Bětnarjec, Fabian Kaulfürst, Maria Maćijowa, Timo Meškank, Anja Šenec, and Michał Wowčer. In addition, Fabian Kaulfürst and Timo Meškank gave me linguistic advice for which I am very grateful. Needless to say, the responsibility for this article is solely mine.

[2] The only generative analysis of Upper Sorbian is that of Rubach (2008), but the discussion is minimal and concerns solely the issue of prepalatal nasals.

complex example of an ONSET-driven conspiracy. Typologically, Upper Sorbian is interesting because its syllable structure is governed by different principles in the domain of the root and in the domain of the word. On the theoretical side, the question is whether Optimality Theory (henceforth OT; Prince & Smolensky [1993] 2004; McCarthy & Prince 1995) which has been designed to solve conspiracies can deal with the complexities of Upper Sorbian. The answer is that it cannot unless it is modified to admit derivational levels. That is, the analysis argues for a derivational model of phonology and against parallel computation. A point of interest is that level 1 in Upper Sorbian must be defined as the bare root level, not as the expected stem level that includes roots and affixes. Further, it is demonstrated that Itô & Mester's (1999) CRISP EDGE constraint makes wrong predictions for Upper Sorbian and *MULTI, a new constraint, needs to be postulated. Also, the analysis bears on the issue of positional markedness VERSUS positional faithfulness and the question of whether, counter to McCarthy (1999), Duke of York derivations should be admitted in phonology.

Upper Sorbian, a minority Slavic language spoken in the eastern part of Germany near the border with Poland, exhibits a complex pattern of disparate processes that are united in their goal to satisfy ONSET. The goal is achieved in several different ways. First, a vowel may turn into a glide, thereby providing an onset to the syllable that would otherwise be onsetless, *ˈjara*[3] 'very', //iara// → [ja.ra].[4] Second, a vowel may glide into the coda, *raj* 'paradise', //rai// → [raj], where gliding precludes the occurrence of *[ra.i] with an onsetless second syllable. Third, post-vocalically, //u// turns into a labial approximant [ʊ], *ˈsawna* 'sauna', //sauna// → [saʊ.na]. Fourth, an onset can be derived by *j*-insertion: *diaˈlekt* 'dialect', //dialɛkt// → [di.ja.lɛkt]. The preference for insertion over gliding in //dialɛkt// is driven by the constraint against complex onsets, banning the candidate *[dja.lɛkt]. But, fifth, this generalization is contradicted by the occurrence of complex onsets in morphologically derived words such as *ˈRomjan* 'inhabitant of Rome', //rom+ian// → [ro.mjan]. Sixth, in some contexts, Upper Sorbian inserts [ʊ] rather than [j], *ˈkanu+wa* 'canoe' GEN.SG., //kanu+a// → [ka.nu.ʊa]. Seventh, *j*-insertion may be spawned not only by high vowels but also by mid vowels, *ˈstereo* 'stereo', //stɛrɛɔ// → [stɛrɛjɔ]. Eighth, mid vowels may also induce the insertion of the approximant, *ˈSamoa* 'Samoa', //samo+a// → [samoʊa]. Ninth, an onset can be provided by initial *ʔ*-insertion, *ˈabo* 'but', //abɔ// → [ʔabɔ] (word-initial *ʔ*-insertion), and, tenth, glottal stop insertion can repair stressed syllables that would otherwise be onsetless (stressed syllable *ʔ*-insertion), *kokaˈin* 'cocaine', //kɔkain// → [kɔkaʔin]. It would appear that this plethora of ten different processes eliminating onsetless syllables would leave no trace of hiatus in

[3] I will adopt the practice of marking the stressed syllable with ˈ. Marking stress is necessary because stress is potentially relevant for some generalizations. Monosyllabic words are always stressed, so there is no need to mark them.

[4] I use double slashes for underlying representations, single slashes for intermediate representations and square brackets for phonetic representations. Also, in accordance with the standard practice, I use dots to mark syllable boundaries. Examples are cited in their orthographic form, with IPA transcriptions added where relevant.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Upper Sorbian phonology, but this is not true: there are words that admit onsetless syllables, as in *ˈdual* 'dual', //dual// → [du.al], *geo ˈgraf* 'geographer', //gɛɔgraf// → [gɛ.ɔ.graf] and *ar ˈchai+sk+i* 'archaic', whose root is //arxai// → [ar.xa.i].

The processes leading to ONSET satisfaction have not been described in the literature on Upper Sorbian before, so this article is the first piece of work stating the relevant generalizations. The data come primarily from *Prawopisny słownik hornjoserbskeje rěče* (Völkel & Meškank 2005), a large dictionary that I have studied in great detail, and from the materials of the Sorbian Language Commission concerning an ongoing debate with regard to the orthographic reforms and the accommodation of borrowings (Maćijowa 2007).

These sources have been complemented by interviews with native speakers that were conducted by me personally during the Sorbian Language and Culture Summer Schools in Bautzen in 2008 and 2014. Summer schools are an excellent opportunity to conduct interviews because all the instructors, lecturers, and the administrative staff are native speakers of Sorbian. During the fieldwork, informants were asked to read word lists. I transcribed what I heard and I made informal recordings so that I could return to the data when needed. The speakers were Sorbian students and instructors in the age bracket 22–50. All of them were born and raised in *Lužica* 'Lausitz', which is the region of Germany where Upper Sorbian still survives as a minority language.

While transcribing and judging the data may be difficult, my task was facilitated in three ways.

First, it helps if the transcriber has the relevant types of data in their native language. This is the situation here: my language has both glides derived from vowels and glottal stops, though the latter are obligatory only under emphasis. For example, I can tell without difficulty if the Upper Sorbian word *ˈklient* 'customer' is pronounced with [i.jɛ], [i.ɛ] or [i.ʔɛ] because my native language has exactly the same word and it is known that the word is pronounced with [i.jɛ], so any other pronunciation would strike me as different from that in my native language.

Second, it also helps if the orthography of a language is closely, or relatively closely, phonetic. Compare the spelling and pronunciation of the same words in Polish and Upper Sorbian.

| (1) | Upper Sorbian | Polish | gloss |
|---|---|---|---|
| | mu*mija* [mija] | mu*mia* [mja] | 'mummy' |
| | Ro*mjan* [mjan] | rzy*mia*nin [mja] | 'Roman' |
| | kan*u* – kan*uw*owy [uʊ][5] | no corresponding ADJ | 'canoe' ADJ |

Polish uses the letter *i* for both [i] and [j] while Upper Sorbian makes the distinction in the spelling.

Third, Upper Sorbian exhibits alternations in which the contrast zero–glide is reflected in the spelling.

[5] I argue later that [ʊ] comes from the glide /w/.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

| (2) | NOM.SG. (zero ending) | GEN.SG. (‘a’ ending) | gloss |
|---|---|---|---|
| | ˈalib*i* [i] | ˈalib*ija* [ij] | ‘alibi’ |
| | ˈem*u* [u] | ˈem*uw*a [uʊ] | ‘emu’ |

All the data used in my research were checked with two professional linguists who are native speakers of Upper Sorbian. Before submission, one of them read an earlier version of this paper to make sure that there are no errors in the data.[6]

This article is organized as follows. Section 1 reviews the basic data. Section 2 provides a preliminary analysis. Section 3 adduces arguments for level distinction. Section 4 studies Upper Sorbian gliding and insertion from the point of view of Stratal/Derivational OT. Section 5 concludes with a summary of the results.

## 1. Data

The goal of this section is to present the data and state the descriptive generalizations. The inventory of surface contrastive segments (phonemes) has been studied in several traditional grammars, including Michalk (1955), Wowčerk (1955), Schuster-Šewc (1968), Stone (1993), and Schaarschmidt (2002). The most recent study is due to Jocz (2011).

Important for this paper is the vowel system, which I cite from Jocz (2011).

(3) Upper Sorbian vowels

| | | |
|---|---|---|
| i | ɨ[7] | u |
| ɪ | | ʊ |
| e | | o |
| ɛ | | ɔ |
| | a | |

The vowel [i] and the glide [j] are in complementary distribution, so I will assume the tenet of autosegmental phonology that [j] is represented as [i] at the melodic tier and is different from the vowel [i] in that it is mora-less and occurs in syllable onsets or codas (Clements & Keyser 1983; Levin 1985; Hayes 1989; and others). To clarify further: underlying representations contain //i// rather than //j// because [j] is predictable and can be derived from //i// by the gliding rules in (4). These rules are perfectly general statements known from a number of languages.

The vowel [i] is the source of [j] in pre-vocalic and post-vocalic contexts, where [j] is an effect of gliding either into the onset (4a) or into the coda (4b).

[6] In my fieldwork, I was not able to look at variability within and across speakers, so this issue awaits further research.

[7] Jocz (2011) assumes that [ɨ] is an allophone of /i/.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(4) (a) Gliding into the onset: i → j / — V[8]

| | |
|---|---|
| ˈjara ‘very’ | ja ‘I’ |
| ˈjama ‘pit’ | ˈjazyk ‘tongue’ |
| ˈjejko ‘egg’ | ˈkajak ‘kayak’ |

(b) Gliding into the coda: i → j / V —

| | |
|---|---|
| raj ‘paradise’ | čaj ‘tea’ |
| ˈspokoj ‘quiet’ | ˈfajfa ‘pipe’ |
| bojˈkot ‘boycott’ | fajn ‘fine’ |

Gliding is inhibited in the morphological root of the word in CiV contexts. ONSET is then satisfied by glide insertion.[9]

(5) *Glide insertion*: Ø → j / Ci — V

| | |
|---|---|
| radiˈator [ra.di.ja.tɔr][10] ‘radiator’ | piˈano [pi.ja.nɔ] ‘piano’ |
| diaˈlekt [di.ja.lɛkt] ‘dialect’ | kliˈent [kli.jɛnt] ‘customer’ |
| diaˈgram [di.ja.gram] ‘diagram’ | biblioˈteka [bi.bli.jɔ.tɛ.ka] ‘library’ |

While gliding into the coda shown in (4b) is the default strategy of resolving a /Vi/ hiatus, it should be noted that there is a small class of words (10 or so) that resist this process.

(6) Exceptionally, no gliding into the coda
arˈchaiski [ʔar.xa.i.ski] ‘archaic’: [a.i], not *[aj]
intuiˈtiwny [ʔin.tu.i.tiʊ.nɨ] ‘intuitive’: [u.i.], not *[uj]
kokaˈin [kɔ.ka.ʔin] ‘cocaine’: /a.i/, not *[aj][11]

The words in (6) are clearly exceptional because the absence of gliding is unpredictable – compare the contrast *raj* [raj] ‘paradise’ but *arˈchaiski* [xa.i-].[12]

As will become clear later, the domain of the word plays an important role in the phonology of Upper Sorbian. The domain is defined as the morphological root plus affixes, an expected definition. The vowel //i// shows two patterns of behavior in this domain. First, it triggers *j*-insertion (7a), a parallel to what we saw in (5). Second, //i// glides to [j] if it is part of an affix (7b).

[8] Gliding into the onset has one exception: it is the morpheme *iˈon* ‘ion’, which is pronounced [ʔijɔn] rather than [jɔn], where the latter follows the pattern in (4a). However, speakers unfamiliar with the word *ion* or with its compounds such as *ionosfera* ‘ionosphere’ tend to regularize the surface form and pronounce it [jɔn].

[9] The orthographic convention is not to reflect the glide in the spelling of the root morpheme.

[10] Here and below, the transcription of palatalization is suppressed as not relevant.

[11] The surface form is [a.ʔi] because [i] is stressed and stressed syllables undergo *ʔ*-insertion. See Section 4.

[12] Here and below, I transcribe the relevant syllables only.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(7) (a) j-insertion in the domain of the word
Ø → j / i — V
ˈrabi ‘rabbi’ NOM.SG. – ˈrab*ij*+*a* [ija] GEN.SG., ˈrab*ij*+*a*t+n+y [ija] ADJ
ˈprofi ‘professional’ NOM.SG. – ˈprof*ij*+*a* [ija] GEN.SG., ˈprof*ij*+*o*w+y [ijɔ] ADJ
ˈkonfeti ‘confetti’ NOM.SG. – ˈkonfet*ij*+*a* [ija] GEN.SG., ˈkonfetij+ow+y [ijɔ] ADJ
ˈtri+o ‘trio’ NOM.SG. – ˈtr*ij*+*a* [ija] GEN.SG., ˈtr*ij*+*o*w+y [ijɔ] ADJ
ˈVerdi ‘Verdi’ NOM.SG. – ˈVerd*ij*+*a* [ija] GEN.SG., ˈVerd*ij*+*o*w+y [ijɔ] ADJ
ˈVisconti ‘Visconti’ NOM.SG. – ˈVisconti*j*+*a* [ija] GEN.SG., ˈVisconti*j*+*o*w+y [ijɔ] ADJ

(b) Gliding into the onset in suffixes
i → j / — V
Rom ‘Rome’ – ˈRom+jan ‘inhabitant of Rome’
dub ‘oak’ – ˈdub+jan+y ADJ
dom ‘house’ – ˈdom+jac+y ADJ

The data in (7b) present a new pattern because gliding creates a complex onset: [ro.mjan], [du.bja.nɨ] and [dɔ.mja.tsɨ]. This is a contrast to the data in both (5) and (7a).

While the gliding in (7b) is a special case, gliding into the coda is the same in the domain of the word as in the domain of the root illustrated in (4b).

(8) Gliding into the coda in the word domain
ˈryb+a ‘fish’ FEM. NOM.SG. – ˈryb+ojt+y ADJ: //rɨb+ɔit+ɨ// → [rɨbɔjtɨ]
ˈkopyt+o ‘hoof’ NEUTER NOM.SG. – ˈkopyt+ojt+y ADJ: //kɔpɨt+ɔit+ɨ// → [kɔpɨtɔjtɨ]
lěs ‘forest’ MASC. NOM.SG. – ˈlěs+ej DAT.SG.: //lɪs+ɛi// → [lɪsɛj]
nan ‘father’ MASC. NOM.SG. – ˈnan+aj NOM. DUAL: //nan+ai// → [nanaj]
ˈstar+y ‘old’ MASC. NOM.SG. – ˈnaj+starš+i ‘oldest’ //nai+starʃʲ+i// → [najstarʃʲi]

The insertion of [j] is spawned not only by the high front vowel //i//, as in (5) and (7a), but also by the mid front vowel, as shown in (9).

(9) j-insertion triggered by a mid front vowel
ˈstere+o [stɛrɛjɔ] ‘stereo’ NOM.SG. – ˈsterej+a GEN.SG., ˈsterej+ow+y[13] ADJ
ˈrode+o [rodɛjɔ] ‘rodeo’ NOM.SG. – ˈrodej+a GEN.SG, ˈrodej+ow+y ADJ
ˈwide+o [ʊidɛjɔ] ‘video’ NOM.SG. – ˈwidej+a GEN.SG, ˈwidej+ow+y ADJ

[13] According to the recommendations of the Sorbian Language Commission (Maćijowa 2007), the representation of [j] in the spelling is optional in the NOM.SG. forms but obligatory in all derived forms. Consequently, ‘stereo’ can be written as either *stereo* or *sterej*+*o* NOM.SG. but *sterej*+*a* GEN.SG. and *sterej*+*ow*+*y* ADJ must be spelled with *j*. The spelling rules have no bearing on the pronunciation: [j] is always pronounced in these words, regardless of whether it occurs in the spelling or not. These rules hold for all words in which *e* is followed by a vowel-initial suffix, as in the examples below.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The pattern shown in (9) is contradicted by vowel clusters in the root-internal position where the same configuration of vowels as in (9) fails to trigger insertion. If the vowel is unstressed, the surface representation exhibits hiatus.[14]

| | | |
|---|---|---|
| (10) | geoˈgraf [gɛ.ɔ.graf] ‘geographer’ | kreaˈtiw+n+y [krɛ.atiʊ.nɨ] ‘creative’ |
| | leoˈpard [lɛ.ɔ.pard] ‘leopard’ | ideaˈlizować [ʔi.dɛ.a.li.zɔ.vaʧ] ‘idealize’ |
| | aˈkordeon[ʔa.kɔr.dɛ.ɔn] ‘accordion’ | ˈneon [nɛ.ɔn] ‘neon’ |

The data using the labial approximant [ʊ] to avoid hiatus belong to two classes of cases. First, [ʊ] comes from //u// and, second, [ʊ] occurs in insertion contexts that parallel those in (7a).

| | | | | |
|---|---|---|---|---|
| (11) | (a) | //saun+a// | → | [saʊna], ˈ*sawna* ‘sauna’ |
| | | //paus+a// | → | [paʊsa], ˈ*pawsa* ‘pause’ |
| | | //autɔr// | → | [ʔaʊtɔr], ˈ*awtor* ‘author’ |
| | | //august// | → | [ʔaʊgust], *awˈgust* ‘August’ |
| | | //aut+ɔ// | → | [ʔaʊtɔ], ˈ*awto* ‘car’ |
| | | //autɔmat// | → | [ʔaʊtɔmat], *awtoˈmat* ‘machine’ |
| | | //aul+a// | → | [ʔaʊla], ˈ*awla* ‘hall’ |

| | | |
|---|---|---|
| (b) | ˈkanu [kanu] ‘canoe’ NOM.SG. | ˈkanuw+a [kanuʊa] GEN.SG., ˈkanuw+ow+y [kanuʊɔʊɨ] ADJ |
| | ˈkengeru [kɛŋgɛru] ‘kangaroo’ NOM.SG. | ˈkengeruw+a [kɛŋgɛruʊa] GEN.SG., ˈkengeruw+ow+y [kɛŋgɛruʊɔʊɨ] ADJ |
| | ˈkakadu [kakadu] ‘cockatoo’ NOM.SG. | ˈkakaduw+a [kakaduʊa] GEN.SG., ˈkakaduw+ow+y [kakaduʊɔʊɨ] ADJ |
| | ˈstatu+a [statuʊa] ‘statue’ NOM.SG. | ˈstatuw+ow+y[15] [statuʊɔʊɨ] ADJ |
| | Nikaˈragu+a [ɲikaraguʊa] ‘Nicaragua’ | nikaˈraguw+ask+i [ɲikaraguʊaski] ADJ |

In parallel to the data in (9), [ʊ] is generated not only by the high back vowel, as in (11), but also by the mid back vowel, as in (12). The process occurs in the domain of the word.

[14] A stressed vowel would trigger *ʔ*-insertion, as I explain later.

[15] As was the case with [j] in (9), words that have a vowel ending in the NOM.SG. have two alternative spellings: with or without *w*, hence *statu+a* or *statuw+a* ‘statue’, *Nikaragu+a* or *Nikaraguw+a* ‘Nicaragua’, *Genu+a* or *Genuw+a* ‘Genoa’, and so forth. However, the pronunciation is always with [ʊ], regardless of how the word is spelled.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(12) ˈbo+a [boʊa] ‘boa’, ˈSamo+a [samoʊa] ‘Samoa’, ˈobo+a [ʔɔboʊa] ‘oboe’

In contexts other than those enumerated in (11) and (12), [ʊ] cannot be generated, so the surface form exhibits hiatus.

(13) ˈdual [du.al] ‘dual’
ˈfebruar [fe.bru.ar] ‘February’,
koaˈlicija [kɔ.a.li.tsi.ja] ‘coalition’,
ˈUeda [ʔu.ɛ.da] (name)

Finally, onsetless syllables can be repaired by glottal stop insertion.[16]

| | | | |
|---|---|---|---|
| (14) | (a) | aˈpryl [ʔaprɨl] ‘April’ | ekˈspert [ʔɛkspɛrt] ‘expert’ |
| | | orˈgan [ʔɔrgan] ‘organ’ | uˈlan [ʔulan] ‘lancer’ |
| | (b) | manuˈal [manuʔal] ‘manual’ | poˈem [pɔʔem] ‘poem’ |
| | | kreˈola [krɛʔola] ‘creole’ | oˈaza [ʔɔʔaza] ‘oasis’ |
| | (c) | ˈabo [ʔabɔ] ‘but’ | ˈoda [ʔoda] ‘ode’ |
| | | ˈale [ʔalɛ] ‘however’ | ˈurna [ʔurna] ‘urn’ |
| | | a [ʔa] ‘and’ | ˈekstra [ʔɛkstra] ‘extra’ |
| | (d) | ˈdual [du.al] ‘dual’ | duˈalita [du.ʔa.li.ta] ‘duality’ |
| | | reaˈlistiski [rɛ.a.li.sti.ski] ‘realistic’ | reˈal [rɛ.ʔal] ‘real’ |
| | | meteoˈrit [mɛ.tɛ.ɔ.rit] ‘meteorite’ | meteˈor [mɛ.tɛ.ʔɔr] ‘meteor’ |

The data in (14a) show that a glottal stop is inserted to provide an onset to word-initial syllables. The same strategy is used when the syllable is stressed (14b). The words in (14c) combine the contexts in (14a) and (14b) because the syllable is both initial and stressed. The alternations in (14d) demonstrate that word-internal *ʔ*-insertion is indeed sensitive to stress: stressed syllables, but not unstressed syllables, have [ʔ].

A reviewer points out that the occurrence of glottal stops is typically conditioned by style, speech rate, and token frequency. They, as well as emphasis and prominence, have been argued to play a role in the use of glottal stops (Schwartz 2012). The data used in this paper have not been investigated for style, speech rate, and register. They all come from the reading of word lists. A point to note is that there was consistency across speakers with regard to glottal stops and glides. Given this consistency, theoretical phonology has the liability to construct a model for interacting generalizations that is able to generate the data produced by the speakers. Facts referring to style, speech rate, frequency, and so forth are of primary concern in usage-based grammars. That is, theoretical modeling and usage-based analyses are independent perspectives in linguistic investigation. This paper looks at Upper Sorbian from a theoretical perspective only, leaving usage-based investigation for future research.

[16] I argue later that these data represent two different generalizations: word-initial insertion and stressed syllable insertion.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

A reviewer points out that glottalization occurs with varying degrees of volume and intensity. Phonetically, glottalization is a continuum ranging from a few irregular glottal pulses to fully fledged glottal stops of varying strengths (Balas 2011; Schwartz 2012; Żygis, Brunner & Moisik 2012). The question therefore is at what point glottalization can be regarded as sufficient to constitute a glottal stop. A larger point here is how gradient phonetic reality can be interpreted as categorical and binary, which is what phonology requires.[17] The answer lies with phonology, not with phonetics. The issue and the solution to the issue are illustrated with a familiar example of how the phonological distinctive feature [±back] is defined by its function in the classification of central vowels.

It is widely agreed on, ever since *The Sound Pattern of English* (Chomsky & Halle 1968), that [±back] is the only feature for making categorical distinctions on the front-back axis in the articulation of vowels. Specifically, the claim is that [±central] does not exist as a distinctive feature in phonology. The consequence is that all vowels must yield themselves to the classification of being either [-back] or [+back]. Front vowels, such as [i], are uncontroversially [-back], back vowels, such as [u], are naturally [+back], but central vowels, such as [ɨ], [ə], and [a], appear to create a problem. The question is where we decide to draw the line on the front-back axis: before central vowels, in which case central vowels would be [+back], or after central vowels, in which case [ɨ] (as well as [ə] and [a]) would be [-back]. The answer is phonological, not phonetic. The argument is that [ɨ] aligns itself with back vowels from the point of view of palatalization. This is exemplified by the presence of palatalization in Russian *stol* [ɬ] 'table', *stol*+*e* [lʲɛ] LOC.SG., sto*l*+*ik* [lʲik] DIMIN. and its absence in *stol*+*u* [ɬu] DAT.SG., *stol*+*om* [ɬɔ] DAT.SG. and *stol*+*y* [ɬɨ] NOM.PL. I conclude that //ɨ// is [+back].

Returning to glottalization in Upper Sorbian, there are two phonological tests that come to mind: palatalization and glide insertion. Descriptively, there are two palatalization rules, $C \rightarrow C^{j} \; / \; - \; V[\text{-}_{back}]$, and $t \; d \rightarrow ʧ^{j} \; ʤ^{j} \; / \; - \; V[\text{-}_{back}]$, as the following data illustrate (Schuster-Šewc 1968).

(15) Sorbian palatalization
  (a) ˈ*di*dl+owa+ć [dʲi] 'fiddle', ˈ*Ti*mo [tʲi] (first name)
  (b) płót [t]'fence' – ˈpłoć+*i*k [ʧʲi] (DIMIN.)

The first rule, $t \; d \rightarrow t^{j} \; d^{j}$, applies in non-derived environments while the second, $t \rightarrow ʧ^{j}$, is restricted to derived environments. The point is that neither of these rules applies across word boundaries, so *płót ˈinternata* 'the fence of the boarding school' is pronounced with [t], not with [tʲ] or [ʧʲ]. The absence of palatalization before word-initial [i] follows automatically if we assume that initial glottalization in ˈ*internata* is a glottal stop because then [t] is not adjacent to [i] in [t ʔi].

[17] The issue of glottalization is complex and requires further study. It is possible that glottalization may not always be complete and that there can be variability within and across speakers.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Another phonological argument for glottal stops is drawn from the application of glide insertion. Historically, Upper Sorbian, but not Polish, had a very productive rule inserting [j] and [w] before word-initial high and mid vowels. Compare the Polish and the corresponding Upper Sorbian data in (16).

(16)

| Polish | Upper Sorbian | gloss |
|---|---|---|
| ił | *ji*ł [ji] | 'mud' |
| ikra | ˈ*ji*kra [ji] | 'spawn' |
| uczyć | ˈ*wu*ćić [ʊu] | 'teach' |
| obraz | ˈ*wo*braz [ʊɔ] | 'picture' |

At some point word-initial glide insertion became inert and borrowings did not develop initial glides.

(17) No initial glides in Upper Sorbian
*i*ˈdol 'idol', *i*mˈport 'import': [ʔi] not [ji]
uniˈwersita 'university', ˈurna 'urn', ˈopera 'opera', [u] and [ɔ], not [ʊu] and [ʊɔ]

The absence of glide insertion in (17) is understandable if we assume that Upper Sorbian changed its strategy of filling word-initial onsets and *ʔ*-insertion was added as a rule. *ʔ*-insertion can do a better job of providing an onset than glide insertion because it is independent of the quality of word-initial vowels, so also words beginning with [a] receive an onset, for example, *ale* [ʔalɛ] 'but'. A reviewer points out that development of *ʔ*-insertion is not surprising for yet another reason: Upper Sorbian is surrounded by languages that have rules of *ʔ*-insertion: German, Czech, and Polish.

In sum, there are two phonological rules that constitute arguments for regarding glottalization on vowels as a glottal stop: palatalization and glide insertion. I conclude that next to gliding, $i \rightarrow j$, and *j*-insertion/*w*-insertion, *ʔ*-insertion is a strategy to provide onsets to onsetless syllables in Upper Sorbian.

## 2. Preliminary analysis

This section provides a preliminary analysis of the data adduced in Section 1. Subsection 2.1 considers simple patterns while Subsection 2.2 debates the status of [ʊ].

### 2.1 *Simple patterns*

The tools of the analysis are the familiar OT constraints (Prince & Smolensky [1993] 2004; McCarthy & Prince 1995) that are stated in a simplified form in (18).

(18)
(a) Onset: Syllables must have onsets
(b) No-Coda: Syllables cannot have codas
(c) *Complex-Onset: No complex onsets



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(d) MAX-Seg: Don't delete a segment
(e) MAX-μ: Don't delete a mora
(f) DEP-Seg: Don't insert a segment, that is, a Root node in terms of Feature Geometry[18]
(g) *ʔ: Don't be a glottal stop
(h) *j: Don't be a high front glide
(i) *[+high]: Don't be [+high]

These constraints are for the most part self-explanatory, but a comment is in order regarding MAX-μ and ONSET. The job of MAX-μ is to penalize gliding because gliding is a process that deletes a mora. This is the tenet of autosegmental phonology which holds that, first, the difference between a glide and a vowel is made in terms of syllable structure and, second, syllable nuclei have a mora (Clements & Keyser 1983; Levin 1985; Hayes 1989; and others). The consequence is that a vowel that turns into a glide loses its mora and syllabifies into the onset or into the coda.[19] The evaluation in (19), which looks at *ˈjara* 'very', //iara// → [jara], makes this point clear. To save space, this evaluation and the others below do not display full syllable structure, focusing on whether the vowel is linked to a mora and hence is a syllable nucleus or whether it is not linked to a mora and hence belongs to the onset or to the coda. Solid lines denote ranking, the right-pointing hand ☞ marks the correct winner.

(19) μ μ μ μ μ
| | | | |
//i a r a// → [i a . r a]

| | ONSET | MAX-μ |
|---|---|---|
| μ μ μ<br>\| \| \|<br>i . a . r a | *!* | |
| ☞ μ μ<br>\| \|<br>i a . r a | | * |

In what follows, I will simplify representations by omitting reference to moras and writing the vowel [i] that is the nucleus of the syllable as [i] and the glide from //i// as [j], that is, [j] will stand for the melodic segment [i] that has lost its mora and hence is a glide.[20]

[18] I assume the Halle-Sagey model of Feature Geometry (Halle 1992; Sagey 1986).
[19] Answering a reviewer's query, let me add that Upper Sorbian has lost length and is weight-insensitive, so codas are never moraic. Moras are underlying because they are not predictable. The reason is that the fleeting vowels of Slavic language, the so-called yers, are represented as mora-less vocalic segments; see Rubach (2016).
[20] A reviewer asks for the reason why the list of the constraints in (18) does not include *HIATUS, a constraint that prohibits vowel clusters: *VV. The problem with *VV is that it is redundant. The



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Returning to the evaluation in (19), generating the correct surface form [ja.ra] requires the ranking: ONSET >> MAX-μ. In (20), I extend the list of output candidates to include [ji.ja.ra], a candidate that violates DEP-Seg, a constraint banning insertion. In order to uphold the result obtained in (19), it is necessary to rank DEP-Seg above MAX-μ, as shown in (20).

(20) //iara// → [ja.ra]

| | ONSET | DEP-Seg | MAX-μ |
|---|---|---|---|
| a. i.a.ra | *!* | | * |
| ☞ b. ja.ra | | | * |
| c. ji.ja.ra | | *!* | |

The ranking of DEP-Seg >> MAX-μ expresses the generalization that gliding is preferred to insertion as a strategy to resolve hiatus.

Gliding into the coda exemplified earlier in (4b) shows that NO-CODA must be ranked below ONSET, //rai// → [raj] *raj* 'paradise'. Furthermore, Upper Sorbian never uses deletion as a strategy to satisfy ONSET, so *//rai// → [ra] is not an option. This generalization is expressed by assuming that MAX-Seg is an undominated constraint, as shown in (21). The absence of ranking is indicated by a broken line.

(21) //rai// → [raj][21]

| | MAX-Seg | ONSET | NO-CODA | MAX-μ |
|---|---|---|---|---|
| a. ra.i | | *! | | |
| ☞ b. raj | | | * | * |
| c. ra | *! | | | |

The generalization in (21) that gliding is preferred to insertion is challenged by the data in (5), such as *dia'lekt* 'dialect', which give preference to insertion over

candidate CV.V violates both *VV and ONSET. The place where *VV and Onset take different paths is the word-initial position. The candidate #V.CV from the input #VCV violates ONSET but not *VV, so the ranking ONSET >> DEP-Seg (with MAX-Seg being undominated) induces insertion as a repair: #VCV → #CV.CV. In contrast the ranking *VV >> DEP-Seg does not make this prediction because the candidate #V.CV does not violate *VV.

[21] Candidates showing insertion are analyzed later in this section.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

gliding: //dialɛkt// → [di.ja.lɛkt]. The difference between ˈ*jara* 'very' in (4) and *dia* ˈ*lekt* 'dialect' is that in the latter but not in the former, gliding would lead to the creation of a complex onset: [ja.ra] VERSUS *[dja.lɛkt]. The undesired candidate *[dja.lɛkt] loses to the desired winner [di.ja.lɛkt] if COMPLEX-Onset is ranked above DEP-Seg, the constraint prohibiting insertion.

(22) //dialɛkt// → [di.ja.lɛkt]

| | ONSET | *COMPLEX-Onset | DEP-Seg | MAX-μ |
|---|---|---|---|---|
| a. di.a.lɛkt | *! | | | |
| b. dja.lɛkt | | *! | | * |
| ☞ c. di.ja.lɛkt | | | * | |

The interaction between DEP-Seg and NO-CODA is illustrated in (23), which looks at //radiatɔr// → [ra.di.ja.tɔr] *radi*ˈ*ator* 'radiator'.

(23) //radiatɔr// → [ra.di.ja.tɔr][22]

| | ONSET | *COMPLEX-Onset | NO-CODA | DEP-Seg | MAX-μ |
|---|---|---|---|---|---|
| a. ra.di.a.tɔr | *! | | * | | |
| b. ra.dja.tɔr | | *! | * | | * |
| c. rad.ja.tɔr | | | **! | | * |
| ☞ d. ra.di.ja.tɔr | | | * | * | |

The evaluation works correctly if DEP-Seg is ranked below NO-CODA as then [ra.di.ja.tɔr] (23d) wins over [rad.ja.tɔr] (23c), the desired result.

The constraint system developed thus far needs further improvement. First, as I show below, the ranking in (23) generates the wrong syllabification for words with an intervocalic consonant cluster. Second, Upper Sorbian admits both *j*-insertion and *ʔ*-insertion (Section 1), so it must be determined why *j*-insertion rather than *ʔ*-insertion applies in (23). The question is relevant since the [a] syllable in *radi*ˈ*ator* is stressed and Upper Sorbian has *ʔ*-insertion in stressed syllables, hence [ra.di.ʔa.tɔr] is certainly a viable contender.

[22] See (26) for the final evaluation of //radiatɔr//.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The ranking *Complex-Onset >> No-Coda in (23) predicts that VCCV should syllabify as VC.CV rather than as V.CCV because it is more important to obey *Complex-Onset than to violate No-Coda. However, the facts are different: Upper Sorbian maximizes onsets, hence *Madrid* 'Madrid', *dobry* 'good', *wotrubny* 'cordial', and *krasny* 'red' are syllabified *Ma.drid, do.bry, wo.tru.bny*, and *kra.sny*, respectively. This pattern can be obtained if No-Coda outranks *Complex-Onset.

(24) //madrid// → [ma.drit][23]

| | No-Coda | *Complex-Onset |
|---|---|---|
| ☞ a. ma.drit | * | * |
| b. mad.rit | **! | |

The second question raised by the evaluation in (23) is why *radiˈator* undergoes *j*-insertion rather than *ʔ*-insertion. As noted earlier, the latter is an option because the [a] syllable in *radiˈator* [ra.di.ja.tɔr] is stressed and Upper Sorbian has a process of *ʔ*-insertion in stressed syllables.

The generalization is that whenever *j*-insertion and *ʔ*-insertion compete over the same string, it is the former that wins over the latter. This generalization is captured by the ranking of the segment inventory constraint banning glottal stops **ʔ* higher than the constraint banning the glide **j*. From the point of view of *radiˈator* in (23), these constraints can be ranked anywhere, for example, as the lowest in the hierarchy, but the evaluation of //rai// → [raj] *raj* 'paradise', with the added candidate containing a glottal stop requires that **ʔ* must outrank No-Coda, as the following tableau documents.

(25) //rai// → [raj][24] Max-[25]

| | Max-Seg | Onset | *ʔ | No-Coda | Dep-Seg | Max-μ | *j |
|---|---|---|---|---|---|---|---|
| a. ra.i | | *! | | | | | |
| ☞ b. raj | | | | * | | * | * |
| c. ra.ʔi | | | *! | | * | | |
| d. ra | *! | | | | | * | |

[23] Upper Sorbian has final devoicing.

[24] The candidate [ra.ji], not considered here, is eliminated by *Multi, a constraint that I discuss later in this section.

[25] The reason why inserting [ʔ] is preferred to the insertion of some other consonant is that *ʔ is the lowest-ranked constraint in the hierarchy of consonant inventory constraints, so inserting [ʔ] is the 'cheapest' option.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

With *ʔ ranked as required and No-Coda >> *Complex-Onset argued for in (24), we need to make sure that these rankings do not have adverse effects on the evaluation of *radiator* 'radiator'. They do not, as (26) shows.

(26) //radiatɔr// → [ra.di.ja.tɔr]

| | Onset | *ʔ | No-Coda | *Complex-Onset | Dep-Seg | Max-μ | *j |
|---|---|---|---|---|---|---|---|
| a. ra.di.a.tɔr | *! | | * | | | | |
| b. ra.dja.tɔr | | | * | *! | | * | * |
| c. rad.ja.tɔr | | | **! | | | * | * |
| ☞ d. ra.di.ja.tɔr | | | * | | * | | * |
| e. ra.di.ʔa.tɔr | | *! | * | | * | | |

To conclude, in situations of conflict, when both *j*-insertion and *ʔ*-insertion are applicable, it is *j*-insertion that must win, so [ra.di.ja.tɔr] must win over [ra.di.ʔa.tɔr].

This generalization is flatly contradicted by the data in (27).

(27) (a) ˈIrka [ʔir.ka] 'Irene'
ˈimpuls [ʔim.puls] 'impulse'
iˈdol [ʔi.dɔl] 'idol'
(b) kokaˈin [kɔ.ka.ʔin] 'cocaine'
heroˈin [hɛ.rɔ.ʔin] 'heroine',
kofeˈin [kɔfɛ.ʔin] 'caffeine'

Given the ranking in (26) and, specifically, the fact that *ʔ outranks *j, we would expect [j], not [ʔ], to fill the onsets in (27). The data are perfectly clear: [j] is never inserted word-initially, see (27a). Neither is it inserted word-medially if the configuration is //Vi//, that is, //Vi// is never turned into [V.ji]. At first glance, it appears that these restrictions can be handled by Distinct Glide, a constraint that bans *ji* syllables (Kawasaki 1982; Rubach 2002). What makes this analysis suspect is the fact that Upper Sorbian freely admits [ji] syllables, for example, *jich* [jix] 'them', *idej+i* 'idea' GEN.SG. The explanation that it is **ji* that blocks glide insertion in (27) collapses when we look at the data in (28).

(28) (a) ˈecho [ʔɛxɔ] 'echo', ˈemu [ʔɛmu], ekoˈnom [ʔɛkɔnɔm] 'steward'
(b) ˈoch [ʔɔx] 'wow', ˈopera [ʔɔpɛra] 'opera', ˈoda [ʔɔda] 'ode'



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

We know from the data in (9) and (12) that mid vowels trigger glide insertion, for example, *ˈstere+o* 'stereo', //stɛrɛ+ɔ// → [stɛrɛjɔ] 'stereo' NOM.SG., and that glide insertion is preferred to glottal stop insertion, so we expect to see initial glides in (28). However, what we find is glottal stops and not glides. Similarly, the input //stɛrɛ+ɔ// should be able to resolve hiatus in two ways: first, by generating [j] from //ɛ//, //stɛrɛ+ɔ// → [stɛrɛjɔ], and, second, by generating [w] from //ɔ/, *//stɛrɛ+ɔ// → /stɛrɛwɔ/,[26] like /w/ (ultimately [ʊ]) is generated in *ˈSamo+a*, //samɔ+a// → /samɔwa/ 'Samoa'. But this is not what we find: *stereo* has [j] and the option of generating [w] is not attested.

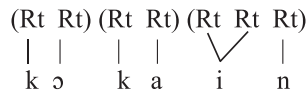
While the data in (27a), (27b), and (28a) as well as (28b) look quite different, they can in fact be reduced to single denominator: the glide cannot be in the same syllable as the vowel that generates it. This is a directionality effect, specifically, the restriction that the glide cannot be to the left of the vowel that spawns it. So, in a $V_1V_2$ configuration, $V_2$ cannot generate a glide. This is exactly what we see in *kokaˈin* 'cocaine': [i] cannot generate a glide because it is a $V_2$ and [a] cannot generate a glide because only high and mid vowels are able to produce glides. Since the hiatus in [ai] cannot be resolved by glide insertion, the grammar moves to the next best option, which is *ʔ*-insertion. This option is available because the *i* of *kokaˈin* is stressed. The *ia* configuration in *radiˈator* is the reverse of the *ai* configuration of *kokaˈin*, but in *ia* the *i* is a $V_1$ vowel, so it is free to spawn a glide: //radiatɔr// → [ra.di.ja.tɔr]. In //stɛrɛ+ɔ// → [stɛrɛjɔ] *stereo*, we see [j] because //ɛ//, the spawning vowel is a $V_1$. In contrast, the //ɔ// of //stɛrɛ+ɔ// is a $V_2$, and, consequently, cannot generate a glide, so **sterewo* is not attested.

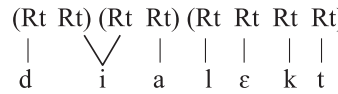
The directionality of glide insertion has been noted as a problem for OT by Itô & Mester (1999). Their solution is to postulate a new constraint called CRISP EDGE: 'multiple linking between prosodic categories is prohibited'. CRISP EDGE bans rightward insertion from $V_1$ in a //$V_1V_2$// cluster, but the facts of Upper Sorbian require exactly the opposite: the insertion from $V_1$ is attested while the insertion from $V_2$ is not. Since the inserted glide must share the features with the spawning vowel, the feature tree of the glide and the vowel is either the same or partly the same. The former occurs if the glide is a full copy of the vowel, as in //Vi//→ [V.ji]. The latter happens when the glide is a partial copy of the vowel, as in //Vɛ// → [V.jɛ]. I illustrate the point in (29), where (29a) shows the undesired candidate *[kɔ.ka.jin] for *kokaˈin* 'cocaine' and (29b) displays the desired winner [di.ja.lɛkt] for *dialekt* 'dialect'. I focus on the relationships between the Root nodes (marked RT) and their melodic content. Syllables are enclosed in parentheses.

[26] The /w/ would surface as [ʊ]; see the discussion in Subsection 2.2.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(29) a. 


b. 


The distinctive property of the banned configuration in (29a) is the occurrence of multiple linking (an effect of spreading) inside one syllable, hence I propose the following constraint.[27]

(30) No-Multiple-Linking (*Multi)
No multiple linking in the domain of one syllable.[28]

Given the ranking *Multi >> *ʔ, it is predicted, correctly, that *koka'in* must be [kɔ.ka.ʔin] because *[kɔ.ka.jin] violates *Multi. A further, beneficial consequence is that word-initial syllables beginning with high and mid vowels cannot spawn glides, so *'Irka* 'Irene' and *'echo* 'echo' do not develop [j]. This is predicted because the putative [j] would have to be in the same syllable as the spawning vowel: [ji] and [jɛ], which is a violation of *Multi. Deprived of their ability to spawn a glide, initial syllables fall prey to the next best option and undergo *ʔ*-insertion, which generates the attested surface forms [ʔir.ka] and [ʔɛxɔ].

As noted in (16), historically *Multi must have been ranked lower than it currently is, which permitted *j*-insertion to create [ji] syllables.[29] This is attested in a class of six words and a few proper names, such as *jich* 'their' and *Jitk* (name), which had [i] rather than [ji] in Old Upper Sorbian (Schuster-Šewc 1983). The current pattern is not to create [ji] syllables, a generalization that is evidenced by *Irka* '*Irene*' in (27) and many similar examples. The [j] in *jich* can still be derived but it must come from gliding rather than from insertion, that is, the underlying representation of *jich* is //iix// rather than //ix//.

[27] The observation that languages may require exactly the reverse of Itô & Mester's (1999) Crisp Edge was first made by Rubach (2002), but Rubach did not state the relevant constraint.

[28] Feature sharing typically occurs across syllables, for example, Nasal Assimilation in *impolite* [ɪm.pə.laɪt]. However, the constraint is certainly violable: voice assimilation in Polish occurs both across syllables and inside one syllable in the Polish: *brat* [d] *Basi* 'Barbara's brother' and *jazd+a* [zd] 'travel' nom.sg. *–jazd* [st].

[29] Similarly, *w*-insertion is responsible for the etymologically epenthetic [ʊ] in words such as *wučer* 'teacher'. As in the case of [ji], the current pattern is not to insert [ʊ] (or rather /w/ spelled out as [ʊ], as I argue later) in words that begin with *u*, hence we have [ʔu] rather than [ʊu] in *uniwersita* 'university' and *urna* 'urn'. The [ʊ] in *wučer* is no longer derivable, but this is not a problem since //ʊ// is an underlying segment in Upper Sorbian, so the initial syllable is simply //ʊu//.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The discussion of *MULTI is summarized by looking at the evaluation of *koka'in* 'cocaine'. Since *MULTI is surface-true, I will assume that it is undominated.

(31) //kɔkain// → [kɔ.ka.ʔin]

| | *MULTI | ONSET | *ʔ | NO-CODA | *COMPLEX-Ons | DEP-Seg | MAX-μ | *j |
|---|---|---|---|---|---|---|---|---|
| a. kɔ.ka.in | | *! | | * | | | | |
| b. kɔ.ka.jin | *! | | | * | | * | | * |
| ☞ c. kɔ.ka.ʔin | | | * | * | | * | | |

The result is correct, but (31) has not considered one important candidate: *[kɔ.kajn] with gliding, *i* → *j*. This candidate would have won in (31) because MAX-μ, the constraint penalizing gliding, is below DEP-Seg, which militates against insertion. The creation of a complex coda in *[kɔkajn] cannot be the explanation here because the GEN.SG. form *koka'in+a* would not have a complex coda and yet the candidate *[kɔ.kaj.na] must lose to the attested surface form [kɔ.ka.ʔi.na]. Similarly, it is irrelevant that the [i] in *koka'in* is stressed because gliding into the coda can be inhibited also when the [i] is unstressed, as in *ar'chaiski* [xa.i] 'archaic', so stress plays no role. Further, notice that the [xa.i] of *ar'chaiski* 'archaic' and the [aj] of *raj* 'paradise' constitute a near minimal pair in that they contrast in the treatment of *i*: //xai// → [xa.i] VERSUS //rai// → [raj].

The default pattern is the one represented by *raj*, that is, gliding is the norm. As observed in Section 1, the absence of gliding is found in a small class of morphemes (10 or so) which are simply exceptions. A further question is how this fact should be encoded in the underlying representation. I follow Rubach (2000a) and assume that the vowel which escapes gliding is prespecified with a sigma, that is, it is prespecified as a syllable nucleus.[30] IDENT-Nuc is then responsible for the blocking of gliding.

(32) IDENT-Nuc
The nucleus on the vowel in the input representation must be preserved on a correspondent of that vowel in the output.

Since IDENT-Nuc is never violated in the surface forms of Upper Sorbian, I will assume that it is undominated and ranked above ONSET.

The evaluation of *koka'in* from (31) is now continued in (33). I mark the prespecified nucleus with N and ignore other aspects of the representation: the reference to moras and the complete syllable trees. Instead, I mark syllable boundaries with a dot. To keep the tableau within reasonable bounds, I omit COMPLEX-Onset because it is not violated by any candidate and hence plays no role.

[30] Like Rubach (2000a; and 2019a), I assume Underspecification rather than the Richness of the Base as a mechanism accounting for predictable information.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(33) N N[31]
| |
//k ɔ k a i n// → [kɔ.ka.ʔin]

| | IDENT-Nuc | *MULTI | ONSET | *ʔ | NO-CODA | DEP-Seg | MAX-μ | *j |
|---|---|---|---|---|---|---|---|---|
| N<br>a. kɔ.ka.in | | | *! | | * | | | |
| b. kɔ.kajn | *! | | | | * | | * | * |
| N<br>c. kɔ.ka.jin | | *! | | | * | * | | * |
| N<br>☞ d. kɔ.ka.ʔin | | | | * | * | * | | |

The evaluation in (33) yields the attested surface form.[32]

Finally, the preliminary analysis of gliding and epenthesis that this section has undertaken needs to account for the preference for *ʔ*-insertion over *j*-insertion in vowel clusters involving mid vowels, specifically //ɛV//. An illustrative example is the word *oceˈan* 'ocean', whose final syllable is stressed, //ɔtsɛan// → [ʔɔtsɛʔan]. The problem is how to exclude the candidate *[ʔɔtsɛjan], which is a viable contender because, as remarked in Section 1 and discussed further in Section 4, mid vowels may spawn glides.

The crucial observation leading to the exclusion of *[ʔɔtsɛjan] is that a glide spawned by a mid vowel is deficient because it cannot draw the feature [+high] from the spawning vowel. That is, if the glide from //ɛ// were a copy of the vowel, we would generate a mid glide rather than a high glide. However, mid glides are prohibited in Upper Sorbian, a generalization that is captured by an undominated constraint on glide well-formedness: glides must be [+high]. In order to obey this constraint, the glide from //ɛ// would have to acquire the feature [+high], which violates the feature markedness constraint stated in (18i): *[+high].[33]

The analysis of *oceˈan* is now straightforward: *[+high] must outrank **ʔ**. In (34), I ignore the first syllable of *oceˈan*, which has a glottal stop, and postpone the discussion of initial *ʔ*-insertion until Section 4.

[31] I add the node Nuc for expository purposes. Technically, Nuc is a vowel that is linked to a mora which in turn is linked to the sigma.

[32] A reviewer points out that IDENT-Nuc is a brute force solution. True, but I wish to mention two things. First, IDENT-Nuc, acting on prespecified nuclei is a way of encoding exceptions and exceptions by their very nature do not succumb to a natural analysis. Second, IDENT-Nuc is motivated independently by cyclic effects that I discuss in Subsection 4.3. See the evaluation of *rabi* 'rabbi' – *rabij*+*a* GEN.SG. in (54).

[33] Needless to say, *[+high] must be outranked by IDENT[+high], which enforces the retention of [+high] on segments that inherit this feature from the input representation.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(34) //tsɛan// → [tsɛʔan]

| | Id-Nuc | *Multi | Onset | *[+high] | *ʔ | No-Coda | Dep-Seg | Max-μ | *j |
|---|---|---|---|---|---|---|---|---|---|
| a. tsɛ.an | | | *! | | | * | | | |
| b. tsɛ.jan | | | | *! | | * | * | | * |
| ☞ c. tsɛ.ʔan | | | | | * | * | * | | |

To conclude, the constraints discussed in this section are listed in (35) that provides a summary of the rankings.

(35) Ranking
Ident-Nuc, Max-Seg, *Multi >> Onset >> *[+high] >> *ʔ >> No-Coda >> *Complex-Onset >> Dep-Seg >> Max-μ, *j[34]

## 2.2 *The status of [ʊ]*

The goal of this section is to present evidence that the bilabial approximant [ʊ] occurring in the processes of gliding and epenthesis is best analyzed as derived from an intermediate /w/ by a process of consonantization: *w* → *v*.[35] I present four arguments in favor of this analysis.

### 2.2.1 *Argument 1*

As shown in Section 1, [ʊ] is derived from //u// when //u// is preceded by a vowel, as in *ˈsawn+a* 'sauna': //saun+a// → [saʊna]. From the structural point of view, this derivation is parallel to the derivation of [j] from //i// in words such as *ˈfajf+a* 'pipe': //faif+a//→ [fajfa].

(36) (a) //faif+a// → [faj.fa]

| | Onset | Max-μ |
|---|---|---|
| a. fa.i.fa | *! | |
| ☞ b. faj.fa | | * |

[34] The constraint *Onset-*w* that I introduce in the following section is ranked above Onset.
[35] In Section 4.3, I explain the characterization of *w* → *v* as consonantization.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(b) //saun+a// → [saw.na]

| | ONSET | MAX-μ |
|---|---|---|
| a. sa.u.na | *! | |
| ☞ b. saw.na | | * |

The structural parallel between (36a) and (36b) holds on the condition that //u// turns into a glide, exactly as //i// turns into [j]. The glide has the melodic representation of the vowel, a standard assumption in autosegmental phonology. A further process turns the /w/ from (36b-ii) into a labial approximant consonant: /w/ → [ʋ]. This simple analysis is possible only if /w/ is admitted as an intermediate representation in the derivation of [ʋ] from //u//.

### 2.2.2 *Argument 2*

A similar argument derives from the observation that //uV// clusters do not change in the sense that //u// is never turned into [ʋ]. For example, the name *Ued+a* //uɛd+a// retains //u// in the surface representation.[36] The absence of [ʋ] is easily accounted for if we assume that //u// would first need to glide to /w/ before ultimately yielding [ʋ] by consonantization. The argument here is that the gliding to /w/ is banned by ONSET-*w*, a constraint that is well known from the study of languages other than Upper Sorbian. For example, ONSET-*w* plays an important role in the analysis of Polish and Slovak (Rubach 2000a).

(37) ONSET-*w*
No [w] in the syllable onset.

All that is required is that ONSET-*w* outrank ONSET, as shown in (38).

(38) //uɛd+a// = [u.ɛ.da] (no change)[37]

| | ONSET -*w* | ONSET | MAX-μ |
|---|---|---|---|
| ☞ a. u.ɛ.da | | ** | |
| b. wɛ.da | *! | | * |
| c. u.wɛ.da | *! | | |

[36] Even though *Ueda* and *UEFA* are probably the only widely used examples of word-initial *uV* clusters, there is no excuse to ignore them. What counts for a theoretical analysis (in contrast to a usage-based analysis) is not the numbers but the prediction of how *uV* structures are pronounced.

[37] Actually, the surface representation of *Ueda* is [ʔuɛda] by word-initial *ʔ*-insertion, which I discuss in Sections 4.2 and 4.3.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Candidate (38c) shows that *Multi and Onset-*w* are independent, as (38c) violates the latter, but not the former, constraint (see also Section 4.2).

### 2.2.3 *Argument 3*

In surface terms, what appears to be *v*-insertion and what we know to be *j*-insertion differ systematically as processes in their operation inside the root morpheme (but not elsewhere, see below). Specifically, the configuration //CiV// spawns a glide but the configuration //CuV// does not.

(39) ˈdiet+a 'allowance': //diɛt+a// → [di.je.ta]
**versus**
ˈdual 'dual': //dual// = [du.al],[38] not *[du.ʊal]

The absence of *[duʊal] is accounted for if we make the assumption that [ʊ] is derived from /w/ by consonantization. On this assumption, the unattested representation *[duʊal] would have to derive from the intermediate representation /duwal/ by *w*-insertion, a parallel to *j*-insertion in *dialekt*. However, the intermediate representation /duwal/ can never be the winner in the evaluation of //dual// because Onset-*w* prohibits /w/ in the onset. In the same vein, the candidates /sawuna/ and /uwɛda/ could never be the successful contenders because they violate Onset-*w*.

### 2.2.4 *Argument 4*

The final argument for intermediate /w/ comes from the directionality of glide insertion. This came up as a problem in the discussion of *kokain* 'cocaine' in (31) and (33). Recall that the constraint system was unable to exclude the candidates containing [ji], so *[kɔ.ka.jin]. The problem is more general than the absence of [ji] where the glide comes from insertion. This is exemplified in (40).

| | | | |
|---|---|---|---|
| (40) | (a) | ˈkanu 'canoe' MASC. NOM.SG.: | //kanu// = [ka.nu] |
| | | ˈkanuw+a GEN.SG.: | //kanu+a// → /ka.nu.wa/ → [ka.nu.ʊa] |
| | | ˈbo+a 'boa' FEM. NOM.SG: | //bo+a// → /bowa/ → [boʊa] |
| | (b) | ˈrabi 'rabbi' MASC. NOM.SG. | //rabi// = [ra.bi] |
| | | ˈrabij+a GEN.SG: | //rabi+a// → [ra.bi.ja] |
| | | ˈrabij+om INSTR.SG: | //rabi+ɔm// → [ra.bi.jɔm] |
| | | | not *//rabi+ɔm// → */ra.bi.wɔm/ → *[ra.bi.ʊɔm] |

As argued in the preceding section, see (31), the absence of *[kɔ.ka.jin] is accounted for by *Multi that outlaws glides occurring in the same syllable as the spawning vowel, as in *[kɔ.ka.jin]. The absence of *[ra.bi.ʊɔm] in (40b) can be accounted for in the same way if we admit an intermediate representation with a

[38] [du.al] is disyllabic, so [ua] is not a diphthong.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

glide: */ra.bi.wɔm/. This candidate is excluded by *MULTI because /w/ and the spawning vowel /ɔ/ are in the same syllable. If the intermediate /w/ does not exist and hence the candidate has the approximant /ʊ/ in /ra.bi.ʊɔm/, the analysis fails: *MULTI is inapplicable because it has jurisdiction over the structure that comes from spreading (multiple linking), not from independent insertion. That is, the candidate /ra.bi.wɔm/ with /w/ from spreading, but not the candidate /ra.bi.ʊɔm/ with /ʊ/ from insertion, is within the purview of *MULTI. I conclude that it is beneficial to admit intermediate /w/ and derive [ʊ] at a later point by consonantization, *w* → *ʊ*.

To conclude, I have adduced four different arguments showing that surface [ʊ] from //u// coming from gliding or from insertion must go through an intermediate stage at which it is the glide /w/. Building on this conclusion, among other things, the following section argues that the correct analysis of Upper Sorbian must admit level distinction.

## 3. LEVEL DISTINCTION

The goal of this section is to adduce evidence for a derivational analysis. Given OT, derivationality is implemented as the framework of Stratal/Derivational Phonology. I argue for an analysis based on derivational levels and present seven different arguments for it, supported by the data from Upper Sorbian.[39]

### 3.1 *Argument 1*

A powerful argument for level distinction was made in the preceding section. The argument is that we need an intermediate stage with the glide /w/ in the derivation of [ʊ] from back vowels by gliding or insertion. For this analysis to work, the consonantization process *w* → *ʊ* must take place at a later level.

### 3.2 *Argument 2*

The domains of morphological roots and words (roots plus affixes) are systematically different for *w*-insertion spawned by //u//: *w*-insertion occurs in the domain of words, but not in the domain of roots.[40]

(41) (a) ˈkanu ‘canoe’ NOM.SG.: //kanu// = [kanu]
ˈkanuw+a GEN.SG: //kanu+a// → /kanuwa/ → [kanuʊa]
**versus**
ˈjanuar GEN.SG: //ianuar// → [januar]

[39] A reviewer draws my attention to the fact that an alternative analysis to enriching computation by introducing levels is to enrich representations. For the latter approach, see Charette (1991) and Van Oostendorp (2007).

[40] While *w*-insertion looks like an example of a Derived Environment (DE) application, other ONSET-sensitive processes (gliding, *j*-insertion and *ʔ*-insertion) are not sensitive to DE. Once we construct a grammar for these other ONSET-filling processes, accounting for *w*-insertion will follow suit. That is, there is no need to bring in DE because the grammar that is already in place generates the facts of *w*-insertion without any additional complications.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(b) ˈkakadu ‘cockatoo’ NOM.SG: //kakadu// = [kakadu]
ˈkakaduw+a GEN.SG: //kakadu+a// → /kakaduwa/ → [kakaduʊa]

**versus**
ˈdual ‘dual’: //dual// = [dual]

Notice that the words in (41) constitute near minimal pairs in the sense that the same vowel configuration //ua// exhibits *w*-insertion if the vowels straddle a morpheme boundary, but not if the vowels are root-internal.

### 3.3 *Argument 3*

Mid vowels spawn glides in the domain of the word (42a), but not in the domain of the root (42b).

(42) (a) ˈstere+o ‘stereo’ NEUTER NOM.SG.: //stɛrɛ+ɔ// → [stɛrɛjɔ]
ˈsterej+a GEN.SG: //stɛrɛ+a// → [stɛrɛja]
ˈsterej+ow+y ADJ: //stɛrɛ+oʊ+ɨ// → [stɛrɛjoʊɨ]
ˈbo+a ‘boa’ FEM. NOM.SG.: //bo+a// → /bowa/ → [boʊa]
ˈbow+ow+y ADJ: //bo+oʊ+ɨ// → /bowoʊɨ/ → [boʊoʊɨ]

(b) geoˈgraf ‘geographer’ MASC. NOM.SG: //gɛɔ-// = [gɛɔ-]
koaliˈcija ‘coalition’ FEM. NOM.SG.: //kɔa-// = [kɔa]

These examples show that the cluster of //ɛ// and //ɔ// spawns a glide in the domain of the word, but not in the domain of the root: //stɛrɛ+oʊ+ɨ// → [stɛrɛjoʊɨ] (42a) versus //gɛɔgraf// = [gɛɔgraf] (42b).

### 3.4 *Argument 4*

The string //CiV// induces glide insertion in roots but not in affixes, which exhibit gliding. The examples in (43) are near minimal pairs.

(43) ˈtymian ‘thyme’: insertion, //tɨmian// → [tɨ.mi.jan]
**versus**
ˈRom+jan ‘inhabitant of Rome’: gliding, //rom+ian// → [ro.mjan]; compare Rom [rom] ‘Rome’
ˈpian+o ‘piano’: insertion, //pian+ɔ// → [pi.ja.nɔ]
**versus**
ˈKamerun+jan ‘inhabitant of Cameroon’: gliding, //kamɛrun+ian// → [kamɛru.njan]; compare ˈKamerun ‘Cameroon’



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

### 3.5 *Argument 5*

Whether the root vowel glides under affixation depends on whether the affix is a suffix or a prefix. There is no gliding when a suffix is added, as in *rabi* 'rabbi' (NOM. SG.) – *rabij+a* (GEN.SG.): [rabija], not *[rabja]. In contrast, the addition of a prefix does not inhibit gliding, as in *ˈz+jednać* 'bring peace': //z+iɛdn+a+ʃʲ// → [zjɛdnaʃʲ], not *[zijɛdnaʃʲ]. That is, [rabija] and [zjɛdnaʃʲ] display different patterns of behavior.

### 3.6 *Argument 6*

As mentioned before, *ʔ*-insertion repairs root-initial onsetless syllables. Root-internal syllables (if unstressed) remain unaffected and exhibit hiatus. This is exemplified by the following near minimal pairs.

(44) *A* ˈmerika [ʔa] 'America' – k*oa*liˈcija [kɔ.a] 'coalition'
*o* ˈperować [ʔɔ] 'do a surgery' – t*eo* ˈlog [ɛɔ] 'theologician'
*u*niˈwersita [ʔu] 'university' – KoloˈS*eu*m [ɛ.u] 'Colosseum'
*I* ˈran [ʔi] 'Iran' – int*ui* ˈtiwny [u.i] 'intuitive'

The problem is how to make sure that *ʔ*-insertion applies root-initially but not root-internally. A similar problem appears when we consider //$V_1V_2$// strings of which $V_2$ is stressed, as I explain below.

### 3.7 *Argument 7*

As noted in (14b–c) in Section 1, a glottal stop is inserted to provide an onset for a stressed syllable. This generalization accounts for the alternations in (45).

(45) ˈdual [dual] 'dual' – duˈalita [duʔalita] 'duality'
reaˈlistiski [rɛalistiski] 'realistic' – reˈal [rɛʔal] 'real'
meteoˈrit [mɛtɛɔrit] 'meteorite' – meteˈor [mɛtɛʔɔr] 'meteor'

The problem is how to guarantee that *ʔ*-insertion applies in stressed syllables and leaves unstressed syllables unscathed.

A heavy-handed analysis of word-initial *ʔ*-insertion in (44) and stressed syllable *ʔ*-insertion in (45) would be to posit new constraints such as those in (46).

(46) (a) ONSET$_{[\text{initial } \sigma]}$
Word-initial syllables must have an onset.
(b) ONSET$_{[\text{stressed } \sigma]}$
Stressed syllables must have an onset.

The problem with these constraints is that they simply state the descriptive facts. An insightful analysis would be one that generates surface representations from an interaction of independent generalizations. Further, ONSET$_{[\text{initial } \sigma]}$ and ONSET$_{[\text{stressed } \sigma]}$



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

are flawed also from a theoretical point of view: they designate two new LOCI for markedness constraints because markedness is now relativized to word-initial syllables and stressed syllables. The effect is that all markedness constraints are tripled in number [*sic*]. Looking at ONSET, for example, we have the generic ONSET stated in (18a) in Section 2.1 and the two specific onset constraints in (46). The theory predicts such tripling for every markedness constraint, a formidable increase in the power of the grammar.[41] This situation is made worse by the fact that we already have such triple editions of faithfulness constraints because, in accordance with the tenets of positional faithfulness, faithfulness is relativized to word-initial syllables and to stressed syllables. That is, an identity constraint represented symbolically as IDENT-X appears in three shapes: IDENT-X (generic), IDENT-$X_{[\text{initial } \sigma]}$ and IDENT-$X_{[\text{stressed } \sigma]}$.

With regard to positional faithfulness, the motivation and the rationale for such distinctions have been argued for in a convincing way, notably by Beckman (1997, 1999) and Casali (1997). I will therefore assume positional faithfulness in the analysis that follows and argue that the addition of the positional markedness constraints in (46) is unnecessary. The argument is built on the assumption that the grammar admits level distinction, an assumption that follows naturally from the seven independent arguments presented earlier in this section.

The theoretical framework of the analysis is that of Stratal/Derivational Optimality Theory (Kiparsky 1997, 2000; Rubach 1997, 2000a,b; Bermúdez-Otero 1999, 2013, 2018; and others). The idea is that evaluation proceeds at three levels or strata: the stem level, the word level, and the post-lexical (post-syntactic) level.[42] The input to the first level is the underlying representation of the stems. The optimal output from the first level is the input to the second level (a new 'underlying representation') and, likewise, the winner candidate from the second level is the input to the third level. Inside a level/stratum, evaluation is fully parallel, as in classic OT. Constraints may be reranked between levels but the reranking must be minimal.

While the three levels are part of the general model of Stratal/Derivational OT, the determination of what constitutes a stem, a word, or a clitic phrase and a sentence is a language-specific matter.

The stem is a general concept: stems are bare roots and roots expanded by affixation; schematically: $[_{\text{Stem}}[_{\text{Stem}}[_{\text{Stem}}[_{\text{Stem}}\ \text{Root}]_{\text{Stem}} + \text{Suffix 1}]_{\text{Stem}} + \text{Suffix 2}]_{\text{Stem}} + \text{Suffix 3}]_{\text{Stem}}$, and so forth. The interest of this paper is that Upper Sorbian requires defining level 1 inputs as bare roots. In a typical situation, level 1 inputs are

[41] A different concern is that once [initial σ] and [stressed σ] are admitted as loci for markedness constraints, nothing in the system can preclude postulating absurd constraints such as *$[+\text{voice}]_{[\text{initial } \sigma]}$ and *$[+\text{voice}]_{[\text{stressed } \sigma]}$ penalizing voiced consonants in initial and stressed syllables, respectively.

[42] An additional level called the clitic level was introduced by Rubach (2011). The clitic level is placed between the word level and the post-lexical level. The analysis of Upper Sorbian proposed here does not need the clitic level, so I will proceed on the assumption that there are three levels in Stratal/Derivational OT.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

larger stems. For example, it has been shown by Bermúdez-Otero (2013) that Spanish does not admit bare roots as level 1 inputs. Level 1 inputs in Polish are roots complete with suffixes. Polish prefixes are analyzed at the word level or at the clitic level (Rubach 2016).

A further question asked by a reviewer is whether WFRs (word formation rules) apply at particular levels, as in classic Lexical Phonology. In this view, English //ɪn//, a classic class 1 prefix, would be added by a WFR at level 1 while English //ʌn//, a classic class 2 prefix, would be added at level 2. This is an attractive assumption but it is potentially problematic because it may lead to affix ordering paradoxes, whereby a level 2 affix might need to be added before a level 1 affix (Kiparsky 1985). This problem goes away if we assume that the default is to do all word formation before phonology. Phonological processing is then guided by the designation of affixes as being level 1 or level 2. The default is that all affixes are processed at level 1. Only designated affixes are processed at level 2. Sometimes the level 2 designation can be predicted by a generalization and hence need not be stipulated. This is the situation in Polish: all prefixes are level 2 (Rubach 2016).

The novelty of Upper Sorbian is that bare roots rather than roots plus affixes are level 1 stems and suffixation comes at level 2. I demonstrate in the subsequent sections that the model of Stratal/Derivational OT is correct and sufficient for an analysis of the complex patterns of generalizations in Upper Sorbian.

## 4. Analysis: Stratal/Derivational OT

This section presents a grammar of Upper Sorbian gliding and epenthesis processes that are active in the Onset conspiracy. Section 4.1 is an overview of where the analysis is heading. Sections 4.2, 4.3, and 4.4 demonstrate how the analysis runs at levels 1, 2, and 3, respectively.

### 4.1 *Overview*

There are three kinds of processes that are active at level 1: gliding, *j*-insertion, and *ʔ*-insertion. The input //i// is glided into the onset or into the coda, depending on whether it is pre-vocalic (gliding into the onset) or post-vocalic (gliding into the coda), as in *jara* 'very', //iara// → [jara], cited in (4a) and *raj* 'paradise', //rai// → [raj], cited in (4b).

The force of gliding is diminished in two ways. First, in a small class of morphemes, underlying //i// is prespecified as a nucleus, the consequence being that it cannot glide because it would offend the undominated Ident-Nuc constraint. For example, *kokaˈin* 'cocaine', with a prespecified nucleus on //i//, cannot claim *[kɔkajn] as the winner (see Section 2). Second, *Complex-Onset thwarts gliding in instances in which it would create a complex onset, so the candidate *[djalɛkt] is not the optimal output from the input //dialɛkt//, *diaˈlekt* 'dialect', as shown in (22) in Section 2. Further, No-Coda makes sure that the input //radiatɔr// *radiˈator*



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

'radiator', cannot evade *Complex-Onset by syllabifying [d] into the coda, so the candidate [rad.ja.tɔr] is doomed, as shown in (26). In sum, words such as *dia'lekt* and *radi'ator* share the generalization that gliding cannot occur with //CiV// inputs. This being the case, //CiV// inputs obey Onset by activating *j*-insertion, so *dia'lekt* and *radi'ator* have [di.ja.lɛkt] and [ra.di.ja.tɔr] as their optimal outputs.

From the point of view of Onset, the hiatus in *dia'lekt* and *radi'ator* could be avoided by either *j*-insertion or *ʔ*-insertion, but the generalization is that Upper Sorbian gives preference to *j*-insertion over *ʔ*-insertion. A glottal stop is inserted only if the glide cannot be inserted. This happens in three situations. First, given that [j] must be spawned by a front vowel, it could not be inserted in roots such as *'Laos* 'Laos' which do not have a front vowel. Second, the vowel *e* in roots such as *geo'graf* 'geographer' cannot generate [j] either because, through the action of *[+high], mid vowels are not permitted to spawn glides at level 1. Third, *Multi prohibits the insertion of [j] before [i], hence *koka'in* 'cocaine' cannot have *[kɔ.ka.jin] as its optimal output. Since *j*-insertion is blocked, the job of filling the onset is passed on to *ʔ*-insertion, hence *'Laos, geo'graf* and *koka'in* have [laʔɔs], [gɛʔɔgraf],[43] and [kɔkaʔin] as their optimal outputs at level 1.

The derivation of /w/ at level 1 is severely constrained: it is limited to gliding into the coda, as in *'sauna* 'sauna' in (36b): //saun// → [sawn].[44] The derivation of /w/ in the onset position is blocked by the undominated Onset-*w*, so *'Ued+a* (name) and *'dual* 'dual' cannot have *[wɛd], *[wuɛd] or *[uwɛd] and *[dwal] or *[duwal] as their optimal outputs. The consequence is that they fall prey to *ʔ*-insertion and leave level 1 with the representations /ʔuʔɛd/ and /duʔal/. In the case of mid vowels, for example, the //ɔ// in *po'et* 'poet', the derivation of [w] is blocked by Onset-*w*: *[pɔ.wet]. The development of [j], however, is thwarted by *Multi. In configurations including //ɛ// as the first vowel of the cluster, as in *oce'an* 'ocean', *j*-insertion is barred from applying by *[+high]. The generalization is that mid vowels cannot spawn glides at level 1. Consequently, the satisfaction of Onset is achieved via *ʔ*-insertion, so *poet* and *ocean* leave level 1 with [pɔʔet] and [ʔɔ.ʦɛ.ʔan] as the winning candidates, which happen to be the attested surface forms.

The ranking of **ʔ* below Onset makes sure that no root can leave level 1 without an onset. The desirable consequence is that all vowel-initial roots pick up a glottal stop, which is exactly what the surface facts of Upper Sorbian require: recall (see Section 1) that the language has an exceptionless process of word-initial *ʔ*-insertion, hence *A'merika* 'America' and *'Indian* 'Indian' have the syllables [ʔa] and [ʔi] in the phonetic representation: [ʔamɛrika] and [ʔindijan].

The resolution of hiatus via *ʔ*-insertion in root-internal vowel clusters delivers the correct result in the cases where the syllable is stressed, such as *po'et* [pɔʔet] and *koka'in* [kɔkaʔin]. In the case of unstressed root-internal syllables such as *'dual*

[43] In the case of *Laos* and *geograf*, the structure derived at level 1 is altered at level 3 by deleting the glottal stop. See below.

[44] The NOM.SG. ending *a* of //saun+a// is not available for evaluation at level 1 because the evaluation is limited to roots.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

'dual', *geoˈgraf* 'geographer' and *arˈchaiski* 'archaic', *ʔ*-insertion overgenerates at level 1, yielding /duʔal/, /gɛʔɔgraf/ and /arxaʔiski/[45] as the winning candidates. The attested phonetic representations are [dual], [gɛɔgraf], and [arxaiski], so the superfluous /ʔ/ must be deleted at a later level. The question is whether this deletion should occur at level 2 or at level 3. There is no doubt that the deletion of /ʔ/ in unstressed syllables is a level 3 operation. The /ʔ/ plays an important role at level 2 because it blocks glide insertion root-internally.

Level 2 has the word as its domain, which means that the evaluation is extended to strings that include roots plus affixes. The level 1 restrictions on gliding, first, no [CjV] outputs and, second, no [w] in the onset, are lifted at level 2. The input /rom+ian/, ˈ*Romjan* 'inhabitant of Rome', has [ro.mjan] as its winning candidate, which is the correct surface representation.[46] Recall that suffixes (here the //ian// of ˈ*Romjan*) are not available at level 1 because Upper Sorbian limits level 1 to roots. Consequently, the input to level 2 is /rom+ian/, where //o//, but not //i//, was syllabified at level 1 and is designated as the nucleus. Therefore, the /i/ of /the suffix /ian/ does not fall within the purview of IDENT-Nuc and is free to glide at level 2: /rom+ian/ → [ro.mjan].

There is no danger that /rabi+a/, the GEN.SG. of ˈ*rabi* 'rabbi', can follow the same path as ˈ*Romjan* at level 2 and undergo gliding. The reason is that *i* → *j* in /rabi+a/ is blocked by the undominated IDENT-Nuc, a constraint that prohibits the gliding of a vowel that has been specified as the nucleus. Given that ONSET is ranked high and **ʔ* remains ranked above DEP-Seg, /rabi+a/ has no option but to undergo *j*-insertion, yielding [rabija], the correct surface form. For the same reason, /kanu+a/, the GEN.SG. of ˈ*kanu* 'canoe', cannot undergo gliding or *ʔ*-insertion. With ONSET-*w* reranked below ONSET at level 2, /kanu+a/ undergoes *w*-insertion and leaves level 2 with the glide /w/: /kanu+a/ → /kanuwa/. The attested surface representation [kanuʊa] is derived at level 3 by consonantization, *w* → *ʊ*.

In contrast to level 1, level 2 is open to the derivation of glides spawned by mid vowels because *[+high] >> **ʔ* is reranked to **ʔ* >> *[+high]. This means that it is better to add the feature [+high], as required when the glide is spawned by a non-high vowel, than to insert [ʔ]. Thus, /stɛrɛ+ɔ/, ˈ*stereo* 'stereo', goes to [stɛrɛjɔ] and /bo+a/ ˈ*boa* 'boa' turns into /bowa/ at level 2 and further to [boʊa] at level 3. The absence of *w*-insertion as well as *j*-insertion in unstressed root-internal vowel clusters containing mid vowels is explained by the fact that these clusters do not exhibit hiatus at level 2. This is so because *ʔ*-insertion overgenerated at level 1, yielding the intermediate representations with a glottal stop: /du.ʔal/, /gɛ.ʔɔ.graf/, and /ar.xa.ʔi.ski/. The clean-up operation deleting /ʔ/ takes place at level 3, where **ʔ* is reranked above ONSET and MAX-Seg.[47]

[45] The //i// in //arxai+sk+i// 'archaic' is prespecified as a nucleus, hence it cannot glide to *[arxajski].

[46] A reviewer asks why ˈ*Romjan* is syllabified [ro.mjan]. I address this issue later: see the derivation in (53).

[47] In Stratal/Derivational OT, there is no distinction between the deletion that affects underlying segments and the deletion of a segment that was inserted on an earlier level.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(47) Level 3
/du.ʔal/ → [du.al]
/gɛ.ʔɔ.graf/ → [gɛ.ɔ.graf]
/ar.xa.ʔi.ski/ → [ar.xa.i.ski]

The deletion of /ʔ/ must not occur in initial syllables and in stressed syllables. This is no problem, however. The positional faithfulness constraints MAX-Seg$_{[\text{initial } \sigma]}$ (no deletion in initial syllables) and MAX-Seg$_{[\text{stressed } \sigma]}$ (no deletion in stressed syllables) are ranked above *ʔ, so /ʔ/ survives the clean-up operation and occurs in the surface representations of words such as *A ˈmerika* [ʔamɛrika] 'America' and *po ˈet* [pɔʔet] 'poet'.[48]

The details of the analysis are presented in the ensuing sections. Section 4.2 examines level 1 evaluations. Section 4.3 lays out the analysis at level 2. Section 4.4 completes the presentation by explaining the changes that occur at level 3.

## 4.2 *Level 1*

The evaluations in Section 2 should be understood now as level 1 evaluations, so most of the phonological processes that are active on level 1 have already been discussed. However, there are two types of cases that require further scrutiny because the formal apparatus was not complete with all the relevant constraints at the time when they were discussed in Section 2. The two cases in point are *j*-insertion originating from mid vowels and the role of *ʔ*-insertion at level 1.

The interaction between glide insertion spawned by mid vowels and *ʔ*-insertion unveils a clear generalization, see (42): mid vowels cannot spawn glides at level 1, so ONSET is satisfied by *ʔ*-insertion. As explained earlier, this generalization is captured by the ranking *[+high] >> *ʔ. The constraint *[+high] prohibits the addition of [+high], an operation that is necessary if a glide comes from a non-high vowel. In (48), I repeat the evaluation from (34) for *oce ˈan* //ɔtsɛan// 'ocean', now extended to include the initial vowel, and add an evaluation for *po ˈet* //pɔet// 'poet'. Irrelevant constraints have been omitted.

(48) (a) *Level 1* //ɔtsɛan// → [ʔɔtsɛʔan]

| | ONSET-*w* | *MULTI | ONSET | *[+high] | *ʔ | NO-CODA | DEP-Seg | *j |
|---|---|---|---|---|---|---|---|---|
| a. ɔtsɛ.an | | | **! | | | * | | |
| b. ʔɔtsɛ.jan | | | | *! | * | * | ** | * |
| c. wɔtsɛ.ʔan | *! | * | | * | * | * | ** | |
| ☞ d. ʔɔtsɛ.ʔan | | | | | ** | * | ** | |

[48] A reviewer remarks that typically glottal stop insertion would be expected to occur post-lexically rather than on level 1. This expectation is true for Lexical Phonology that is guided by the principle of Structure Preservation (Kiparsky 1985). Stratal/Derivational OT has abandoned Structure Preservation (Bermúdez-Otero 2018), so nothing stands in the way of inserting glottal stops on level 1.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(b) *Level 1* //pɔet// → [pɔ.ʔet]

| | ONSET-*w* | *MULTI | ONSET | *[+high] | *ʔ | NO-CODA | DEP-Seg | *j |
|---|---|---|---|---|---|---|---|---|
| a. pɔ.et | | | *! | | | * | | |
| b. pɔ.jet | | *! | | * | | * | * | * |
| c. pɔ.wet | *! | | | * | | * | * | |
| ☞ d. pɔ.ʔet | | | | | * | * | * | |

The words in (48) are the attested surface forms because *ʔ*-insertion has supplied [ʔ] to the root-initial syllable and to the stressed syllables, which is where [ʔ] is found in the surface representations.

The situation is different when root-internal syllables are unstressed. This happens in words such as *ar*ˈ*chaiski* 'archaic', ˈ*du*al 'dual', and *geo*ˈ*graf* 'geographer'. The grammar of level 1 predicts that these syllables will obtain a glottal stop, as shown in (49). The outputs are the intermediate representations, which I enclose in single slashes. Irrelevant constraints that are not violated by any of the candidates or are violated in exactly the same way by all candidates have been omitted. The evaluation in (49a) considers the relevant part of the word *ar*ˈ*chaiski* 'archaic'. Recall from the discussion in Section 2 that the //i// in *archaiski* is prespecified as a nucleus, so IDENT-Nuc excludes the candidate with gliding.

(49) (a) N N[49]
| |
*Level 1* //xai// → /xa.ʔi/

| | ID-Nuc | ONSET-*w* | *MULTI | ONSET | *[+high] | *ʔ | NO-CODA | DEP-Seg | *j |
|---|---|---|---|---|---|---|---|---|---|
| N \| a. xa.i | | | | *! | * | | | | |
| b. xaj | *! | | | | * | | * | | * |
| N \| c. xa.ji | | | *! | | * | | | * | * |
| N \| ☞ d. xa.ʔi | | | | | * | * | | * | |

[49] Here and below, the output is in slashes rather than in square brackets because /xa.ʔi/ is an intermediate representation. The surface forms are derived by *ʔ*-deletion at level 3: /xa.ʔi/ → [xa.i], /du.ʔal/ → [du.al] and /gɛ.ʔɔ.graf/ → [gɛ.ɔ.graf].



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(b) *Level 1* //dual// → /du.ʔal/

| | ONSET-*w* | *MULTI | ONSET | *[+high] | *ʔ | NO-CODA | DEP-Seg | *j |
|---|---|---|---|---|---|---|---|---|
| a. du.al | | | *! | * | | * | | |
| b. du.wal | *! | | | * | | * | * | |
| ☞ c. du.ʔal | | | | * | * | * | * | |

(c) *Level 1* //gɛɔgraf// → /gɛ.ʔɔ.graf/

| | ONSET-*w* | *MULTI | ONSET | *[+high] | *ʔ | NO-CODA | DEP-Seg | *j |
|---|---|---|---|---|---|---|---|---|
| a. gɛ.ɔ.graf | | | *! | | | * | | |
| b. gɛ.wɔ.graf | *! | * | | * | | * | * | |
| c. gɛ.jɔ.graf | | | | *! | | * | * | * |
| ☞ d. gɛ.ʔɔ.graf | | | | | * | * | * | |

To clarify, the constraint *[+high] is violated by any occurrence of [+high] (see Note 33). In candidates (49a-iii) and (49b-ii), it is violated once, not twice, because the glides come from spreading, so they share the feature tree with the spawning vowel.

Glottal stops in unstressed syllables in (49) are deleted at level 3 (see Section 4.4) because the attested surface forms exhibit hiatus: [xa.i], [du.al], and [gɛ.ɔ.graf]. The presence of a glottal stop at the intermediate stage, specifically at the output of level 1, is a Duke of York situation (Pullum 1976). This is not a problem for two reasons. First, as documented in Rubach (2003, 2014, 2019b), OT must admit Duke of York derivations and, second, the intermediate representations with a glottal stop play an important role at level 2 because they account for the absence of glide insertion root-internally. Paradoxically then, the Duke of York derivation here is an asset rather than a drawback. I clarify this reasoning further in the following section. At this point, I conclude that all winning outputs from level 1 have an onset. In most cases, the onset generated at level 1 occurs in the attested surface representation. In some cases, the onset (invariably a glottal stop) exists only in the intermediate representations that are processed further at level 3.

A reviewer asks for an independent example that would motivate Duke of York derivations in OT. A clear example is found in Polish (Rubach 2003). Polish soft labials (underlying or derived by palatalization) are decomposed into a labial and a glide, as in //karp$^{j}$+a// → /karpja/ → [karp$^{j}$j+a] 'carp' GEN.SG. The decomposition process is driven by the segment inventory constraint *SOFT-Labial. Faithfulness to the soft //p$^{j}$// is satisfied by breaking up the input into two output segments: a hard /p/ and a palatal glide /j/, where both are correspondents of //p$^{j}$//.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The correspondents collectively preserve the properties of the input, with [-back, +high] now located on /j/ rather than on the labial. An independent post-lexical process of surface palatalization applying to all consonants before /i/ and /j/ repalatalizes the labial: //p$^{j}$// → /pj/ → [p$^{j}$j/. Thus, soft *p$^{j}$* becomes hard /p/ only in order to revert to soft [p$^{j}$] in the surface representation: a classic Duke of York derivation.[50]

### 4.3 *Level 2*

Level 2 is the domain of the word, which means that structures involving roots and affixes are within the purview of level 2 phonology. In this section, I review the processes which operate in a different way on level 2 than on level 1. The constraints are exactly the same but their ranking may be different.

In some cases, the processes on level 2 are exactly the same as on level 1. Gliding into the coda is a case in point.

(50)

| | |
|---|---|
| nan 'father' | ˈnan+aj masc. NOM.DUAL |
| | ˈnan+omaj DAT./LOC.DUAL |
| ˈryb+a 'fish' FEM. NOM.SG. | ˈryb+ojt+y ADJ MASC. NOM.SG. |
| ˈkopyt+o 'hoof' NEUTER NOM.SG. | ˈkopyt+ojt+y ADJ MASC. NOM.SG. |

The constraint hierarchy from level 1 delivers the correct result, as shown by the evaluation of ˈ*nan+aj* /nan+ai/ 'father' NOM.DUAL in (51). Irrelevant constraints have been omitted.

(51) *Level 2* /nan.+ai/[51] → [na.naj]

| | *Multi | Onset | *ʔ | No-Coda | Complex-Onset | Dep-Seg | Max-μ | *j |
|---|---|---|---|---|---|---|---|---|
| a. na.na.i | | *! | | | | | | |
| ☞ b. na.naj | | | | * | | | * | * |
| c. na.na.ji | *! | | | | | * | | * |
| d. na.na.ʔi | | | *! | | | * | | |

[50] For an OT analysis of this derivation, see Rubach (2003).

[51] The root *nan*, unlike the suffix //ai//, has been processed at level 1, so it enters level 2 with the syllable structure assigned at level 1. The input in (51) is therefore the string /(nan)$_{\sigma}$+ai/. The suffix //ai// is first considered at level 2, so it has no syllable structure in the input. The standard constraints Onset, No-Coda, and *Complex-Onset take care of resyllabification wherever warranted. In the instance at hand, the syllable-final /n/ of /(nan)$_{\sigma}$/ is syllabified as the onset of the newly created syllable *aj*: /(nan)$_{\sigma}$/ → [(na)$_{\sigma}$/(naj)$_{\sigma}$]. In what follows, I will ignore resyllabification because it does not bear on the theoretical issues under discussion.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Gliding into complex onsets highlights a point of distinction between levels 1 and 2 (see the data in (7b) in Section 1). Specifically, hiatus in CiV strings is resolved by insertion at level 1 but by gliding at level 2: //dialɛkt// → [di.ja.lɛkt] at level 1 (see (22) in Section 2) VERSUS /rom.+ian/ → [ro.mjan] ˈ*Romjan* 'inhabitant of Rome' at level 2. This change of strategy is expressed as the reranking of DEP-Seg above NO-CODA.

(52) Reranking at level 2
Level 1: COMPLEX-Onset >> DEP-Seg
Level 2: DEP-Seg >> COMPLEX-Onset

It is this reranking that accounts for the syllabification r*adiˈator* 'radiator', //radiatɔr// → [ra.di.jatɔr] at level 1 VERSUS ˈ*Rom+jan* 'inhabitant of Rome', /rom.ian/ → [ro.mjan] at level 2. To put it simply: due to the low ranking of DEP-Seg at level 1, it is 'cheaper' to insert [j] than to have a complex onset, so //dia// → [di.ja] is better than //dia// → *[dja] in r*adiˈator*. At level 2, DEP-Seg is reranked high, which thwarts glide insertion. The next best option is to have a complex onset because NO-CODA is ranked higher than *COMPLEX-Onset, so /mia/ → [mja] is better than /mia/ → *[mi.ja] in ˈ*Rom+jan*. Importantly, ˈ*Rom+jan* as a word is not available on level 1 because level 1 is a root level, so all that is available is the root *Rom*.

In sum, the evaluation of ˈ*Rom+jan* 'inhabitant of Rome' at level 2 is as follows.

(53) *Level 2* /rom+ian/ → [ro.mjan]

| | ONSET | *ʔ | DEP-Seg | NO-CODA | *COMPLEX-Onset | MAX-μ | *j |
|---|---|---|---|---|---|---|---|
| a. ro.mi.an | *! | | | * | | | |
| ☞ b. ro.mjan | | | | * | * | * | * |
| c. rom.jan | | | | **! | | | * |
| d. ro.mi.jan | | | *! | * | | | * |
| e. ro.mi.ʔan | | * | *! | * | | | |

The level 2 preference for gliding, /CiV/ → [CjV] as in (53), appears to create a problem for the data cited in (7) in Section 1, such as ˈ*rabi* 'rabbi' NOM.SG. – ˈ*rabij+a* GEN.SG. and ˈ*profi* 'professional' NOM.SG. – ˈ*profij+a* GEN.SG. – ˈ*profij+ow+y* ADJ NOM.SG. The problem is how to exclude the undesired candidate *[rabja] from the input /rabi+a/ without losing the insight that ˈ*Romjan* uses gliding as a hiatus resolution strategy: /rom+ian/ → [ro.mjan] in (53).

The solution to this dilemma is already in place and does not require changes in the constraint ranking beyond the DEP-Seg reranking in (52) that is necessary



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

independently. The examples under discussion exhibit a structural difference and that is the key to the problem. The ˈ*rabi* part of the level 2 input /rabi+a/ is a root, so it was processed at level 1 where it was syllabified as /ra.bi/. Crucially, the /i/ is a nucleus when it enters level 2. The constraint IDENT-Nuc, which mandates the retention of the nucleus in the output, thwarts gliding at level 2 in exactly the same way as it thwarts gliding of prespecified inputs at level 1, such as *koka*ˈ*in* 'cocaine', as in (33) in Section 2. This constraint is undominated in Upper Sorbian and, consequently, ranked above ONSET at all levels. The evaluation of /rabi+a/, ˈ*rabij+a*, the GEN.SG. of ˈ*rabi* 'rabbi' proceeds as follows.

(54)


| | IDENT-Nuc | ONSET | *ʔ | DEP-Seg | NO-CODA | *COMPLEX-Ons | MAX-μ | *j |
|---|---|---|---|---|---|---|---|---|
| N N N<br>a. ra.bi.a | | *! | | | | | | |
| N N<br>b. ra.bja | *! | | | | | * | * | * |
| N N N<br>c. ra.bi.ʔa | | | *! | * | | | | |
| N N N<br>☞d. ra.bi.ja | | | | * | | | | * |

In contrast to the //i// in ˈ*rabij+a*, the /i/ of //ian// in ˈ*Rom+jan* was not available at level 1 because it is part of the suffix, not of the root, and suffixes are processed at level 2 because level 1 is the bare root level. Consequently, it was not syllabified as a nucleus and hence escapes the jurisdiction of IDENT-Nuc at level 2. Without the protection of IDENT-Nuc, the /i/ of /ian/ falls prey to gliding, exactly as presented in (53). The distinction between ˈ*rabij+a* and ˈ*Rom+jan* is a classic cyclic effect: we need to process the internal constituent (here the root //rabi//) before we process the external constituent (here the suffixed structure //rabi+a//).

The analysis of //kanu+a// → [kanuʊa] 'canoe' (GEN.SG.) mirrors that of /ra.bi+a/ → [ra.bi.ja] 'rabbi' GEN.SG., but there is one difference: *w*-insertion is no longer blocked by ONSET-*w* due to the following reranking.

(55) Reranking at level 2
Level 1: ONSET-*w* >> ONSET
Level 2: ONSET >> ONSET-*w*



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Level 2 phonology tells a new story about mid vowels. From the point of view of glide insertion, mid vowels were inactive at level 1 because the high-ranking of *[+high] prohibits the addition of [+high], leaving spreading from an existing [+high] on high vowels as the only option. Consequently, clusters such as //ɛɔ// in *geoˈgraf* 'geographer' resolve their hiatus at level 1 by inserting [ʔ] rather than [j]: //gɛɔgraf// → /gɛʔɔgraf/. The same clusters derived at level 2 by affixation spawn [j], which is what we find in the paradigm for ˈ*stere+o* [stɛrɛjɔ] 'stereo' NOM.SG., ˈ*sterej+a* [stɛrɛja] GEN.SG., ˈ*sterej+om* [stɛrɛjɔm] INSTR.SG., and so forth. In terms of the constraint system, this means that the ranking *[+high] >> **ʔ** from level 1 has changed to ***ʔ** >> *[+high] at level 2, so adding [+high] is now better than inserting a glottal stop. Glide insertion does not affect /gɛʔɔgraf/ at level 2 because there is no hiatus. The glottal stop is ultimately deleted at level 3, yielding the attested surface form [gɛɔgraf]. The hiatus produced by **ʔ**-deletion is not repaired because at level 3 ONSET is reranked below DEP-Seg, thwarting insertion.

As noted, relevant for the derivation of *stere+o* 'stereo' (nom.sg.) is the reranking in (56).

(56) Reranking at level 2
Level 1: *[+high] >> *ʔ
Level 2: *ʔ >> *[+high]

The evaluation of ˈ*stere+o* NOM.SG. provides evidence for this change. Recall from Section 2 that glides from mid vowels must add the feature [+high] in order to comply with the requirement that glides must be high. The addition of [+high] is penalized by *[+high]. Irrelevant constraints have been omitted.

(57) *Level 2* /stɛrɛ+ɔ/ → [stɛrɛjɔ]

| | *MULTI | ONSET | *ʔ | *[+high] | DEP-Seg | *j |
|---|---|---|---|---|---|---|
| a. stɛ.rɛ.ɔ | | *! | | | | |
| b. stɛ.rɛ.wɔ | *! | | | * | * | |
| c. stɛ.rɛ.ʔɔ | | | *! | | * | |
| ☞ d. stɛ.rɛ.jɔ | | | | * | * | * |

The absence of *j*-insertion in words such as *reˈal* 'real' and *oceˈan* 'ocean' is accounted for as follows. Words containing vowel clusters undergo **ʔ**-insertion at level 1, generating //rɛal// → [rɛ.ʔal] and //ɔʦɛan// → [ʔɔ.ʦɛ.ʔan]. These are the attested surface forms of *reˈal* 'real' and *oceˈan* 'ocean'. In contrast, ˈ*stere+o* did not have a vowel cluster at level 1 because the root is //stɛrɛ//, so the race was won by the faithful candidate /stɛ.rɛ/ which does not violate ONSET. The hiatus problem first



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

occurs at level 2, at which suffixes enter into the game: the *o* //ɔ// of the NOM.SG. suffix creates a hiatus: /stɛrɛ+ɔ/. But now the constraint ranking is *ʔ >> *[+high], so *j*-insertion is preferred to *ʔ*-insertion, as shown in (57).

The question is how words such as *geoˈgraf* [gɛ.ɔ.graf] whose surface forms exhibit hiatus escape *j*-insertion at level 2. The solution to this problem lies with the overgenerating power of *ʔ*-insertion at level 1. Recall that [ʔ] is inserted by default if a syllable has no onset (see (49c) earlier in this section). Consequently, *geoˈgraf* emerges from level 1 as /gɛ.ʔɔ.graf/. Ultimately, the /ʔ/ must be deleted, but if we delay the deletion till level 3, we have an answer to the question of why the *eo* in *geoˈgraf* does not generate a glide: ONSET is not violated because *geoˈgraf* retains /ʔ/ at level 2. It appears that this analysis does not really explain the absence of *j*-insertion because the problem reemerges at level 3, at which /ʔ/ is deleted and [ɛɔ] constitutes a hiatus. The problem is apparent. I argue in Subsection 4.4 that level 3 phonology does not admit any insertion at all, so the hiatus in *geoˈgraf* is tolerated.

Level 2 bears witness to *w*-insertion, a process that was prohibited at level 1 by the high-ranking ONSET-*w*. The facts and the analysis are parallel to those presented for *j*-insertion spawned by mid vowels and by /i/ in *stere+o* NOM.SG. and *rabij+a* GEN.SG.

We see *w*-insertion in words such as *kanu* 'canoe' NOM.SG. – *kanuw+a* GEN.SG., exemplified in (11b) in Section 1. Relevant here is the GEN.SG. form. Its representation at the input to level 2 is /ka.nu.+a/, where /ka.nu/ is the winner from level 1 and *a* is the GEN.SG. ending first processed on level 2. The vowel cluster in /ka.nu.+a/ triggers *w*-insertion, yielding /ka.nu.wa/. This is a different strategy from that exhibited at level 1, where the same vowel cluster //ua// triggered *ʔ*-insertion, as in *manuˈal* 'manual': //manual// → [ma.nu.ʔal]. The change of the strategy—*ʔ*-insertion at level 1 but *w*-insertion at level 2—is expressed as the reranking of ONSET-*w* >> *ʔ at level 1 to *ʔ >> ONSET-*w* at level 2.

(58) Reranking at level 2
Level 1: ONSET-*w* >> *ʔ
Level 2: *ʔ >> ONSET-*w*

The reason why /kanu+a/ did not receive a glottal stop at level 1 is the same as in the case of /rabi+a/ and /stɛrɛ+ɔ/ analyzed earlier in this section: at level 1 the evaluation is limited to roots and the root //kanu// does not violate ONSET. The violation of ONSET becomes an issue at level 2, at which suffixes, here the GEN.SG. *a*, are within the purview of the constraint system. But at level 2, the ranking is *ʔ >> ONSET-*w*, so *w*-insertion rather than *ʔ*-insertion is the strategy for hiatus resolution. The details of this analysis are shown in (59). Recall from the evaluation of /rabi+a/ in (54) that IDENT-Nuc bans the candidate exhibiting gliding. I omit *MULTI because it is irrelevant.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(59) *Level 2* /ka.nu.+a/ → /ka.nu.wa/

| | IDENT-Nuc | ONSET | *ʔ | *[+high] | DEP-Seg | ONSET-*w* |
|---|---|---|---|---|---|---|
| a. ka.nu.a | | *! | | * | | |
| ☞ b. ka.nu.wa | | | | * | * | * |
| c. kan.wa | *! | | | * | | * |
| d. ka.nu.ʔa | | | *! | * | * | |

The final question is how to avoid *w*-insertion in words that exhibit an [ua] hiatus in the surface representation, such as ˈ*dual* 'dual' NOM.SG. The answer here is the same as with the absence of *j*-insertion in *geo*ˈ*graf*: ˈ*dual* enters level 2 with a glottal stop due to the overgenerating power of *ʔ*-insertion at level 1 (see the evaluation in (49b) in Section 4.2). The input to level 2, /du.ʔal/, has its faithful output /du.ʔal/[52] as the winner because /du.ʔal/ does not violate ONSET and hence there is no incentive to make changes to this representation. At level 3, /du.ʔal/ loses its glottal stop but then all insertion is prohibited, so [du.al] is the predicted surface representation, the correct result.

The analysis of *w*-insertion triggered by mid vowels, as in ˈ*bo+a* 'boa', /bo+a/ → /bo.wa/, uses the mechanisms and the arguments familiar from the analyses of *stere+o* and *kanuw+a*. Examples of roots with *ʔ*-insertion contrasting with *w*-insertion in ˈ*boa* are also parallel: in *kre*ˈ*ola* 'Creole', the glottal stop is retained in the surface representation [krɛ.ʔɔ.la] because the syllable is stressed whereas in *meteo*ˈ*rit* 'meteorite', where *eo* is unstressed, the representation with a glottal stop /mɛ.tɛ.ʔɔ.rit/ is repaired at level 3 to yield the attested surface form [mɛ.tɛ.ɔ.rit].

The evaluation of ˈ*bo+a* on (60) shows the details of the analysis. Irrelevant constraints have been omitted.

(60) *Level 2* /bɔ+a/ → /bo.wa/

| | ONSET | DEP-Seg | *ʔ | *[+high] | ONSET-*w* |
|---|---|---|---|---|---|
| a. bo.a | *! | | | | |
| ☞ b. bo.wa | | * | | * | * |
| c. bo.ʔa | | * | *! | | |

[52] Recall that the slashes indicate intermediate representations. The final output [dual] is derived at level 3 by *ʔ*-deletion.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The winners with /w/ in (59) and (60) are processed further at level 3, where /w/ undergoes consonantization: *w* → *v*.

A reviewer points out that the feature characterization of approximants is controversial. In my analysis, [w] is represented as [u] on the melodic tier, but the [u] is not linked to a mora and hence is not a nucleus. The derivation *w* → *v* occurs on the melodic tier only. Since the melodic segment [u] is vocalic and [ʋ] is not, I assume that the change is from [-cons] to [+cons] and hence I call this process consonantization.

### 4.4 *Level 3*

Level 3 phonology undertakes two actions. First, it repairs the inputs that have a glottal stop in unstressed syllables by deleting the glottal stop and, second, it turns the glide /w/ into an approximant by consonantization: *w* → *v*.

The glottal stops inherited from level 2 are deleted in unstressed syllables in (61a) but retained in stressed syllables in (61b) and word-initially in (61c).

(61) (a) arˈchaiski: 'archaic' /xa.ʔi/ → [xa.i]
ˈdual 'dual': /du.ʔal/ → [du.al]
geoˈgraf 'geographer': /gɛ.ʔɔ.graf/ → [gɛ.ɔ.graf]
meteoˈrit 'meteorite': /mɛ.tɛ.ʔɔ.rit/ → [mɛ.tɛ.ɔ.rit]

(b) kokaˈin 'cocaine': /kɔ.ka.ʔin/ = [kɔ.ka.ʔin]
duˈalita 'duality': /du.ʔa.li.ta/ = [du.ʔa.li.ta]
reˈal 'real': /rɛ.ʔal/ = [rɛ.ʔal]
meteˈor 'meteor': /mɛ.tɛ.ʔɔr/ = [mɛ.tɛ.ʔɔr]

(c) aˈpryl 'April': /ʔa.prɨl/ = [ʔa.prɨl]
orˈgan 'organ': /ʔɔr.gan/ = [ʔɔr.gan]
ekˈspert 'expert': /ʔɛk.spɛrt/ = [ʔɛk.spɛrt]
uˈlan 'lancer': /ʔu.lan/ = [ʔu.lan]

The deletion in (61a) requires that **ʔ* must be reranked above MAX-Seg.

(62) Reranking at level 3
Level 2: MAX-Seg >> *ʔ
Level 3: *ʔ >> MAX-Seg

Given the reranking, /xa.ʔi/ → [xa.i] in *arˈchaiski* 'archaic' is optimal, which is the correct result. The /i/ does not glide to [j] because IDENT-Nuc is undominated at all levels, so it outranks ONSET at level 3.

Finally, the words that lose their glottal stop cannot be permitted to satisfy ONSET by inserting some other segment, for instance, [w] in *ˈdual* and [j] in *geoˈgraf*. That is, the candidates *[du.wal] and *[gɛ.jɔ.graf] must lose to the attested surface forms [du.al] and [gɛ.ɔ.graf]. This result is obtained if DEP-Seg is reranked above ONSET, thwarting all insertion.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(63) Reranking at level 3
Level 2: ONSET >> DEP-Seg
Level 3: DEP-Seg >> ONSET

The evaluations at level 3 for *ar*ˈ*chaiski,* ˈ*dual* and *geo*ˈ*graf* are displayed in (64). Since, as just said, level 3 does not admit glottal stops, gliding or insertion, I will assume at this point that the constraints **ʔ* and DEP-Seg are undominated (but see (66), where I modify this claim).

(64) (a) *Level 3* /xa.ʔi/ → [xa.i] of *ar*ˈ*chaiski*

| | IDENT-Nuc | DEP-Seg | *ʔ | MAX-Seg | ONSET | ONSET-*w* |
|---|---|---|---|---|---|---|
| a. xa.ʔi | | | *! | | | |
| ☞ b. xa.i | | | | * | * | |
| c. xa.ji | | *! | | | | |
| d. xaj | *! | | | | | |

(b) *Level 3* /du.ʔal/ → [du.al]

| | IDENT-Nuc | DEP-Seg | *ʔ | MAX-Seg | ONSET | ONSET-*w* |
|---|---|---|---|---|---|---|
| a. du.ʔal | | | *! | | | |
| ☞ b. du.al | | | | * | * | |
| c. du.wal | | *! | | | | * |
| d. dwal | *! | | | | | * |

(c) *Level 3* /gɛ.ʔɔ.graf/ → [gɛ.ɔ.graf][53]

| | IDENT-Nuc | DEP-Seg | *ʔ | MAX-Seg | ONSET | ONSET-*w* |
|---|---|---|---|---|---|---|
| a. gɛ.ʔɔ.graf | | | *! | | | |
| ☞ b. gɛ.ɔ.graf | | | | * | * | |
| c. gɛ.jɔ.graf | | *! | | | | |

[53] Evaluation (c) is not a case of reranking. These constraints are ranked in the same way at all levels, but the earlier analysis has not disclosed this ranking.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

The retention of glottal stops in the surface representations (61b–c), such as [kɔ. ka.ʔin] and [ʔa.prɨl] is an effect of the following constraints.

(65) (a) MAX-Seg$_{[\text{stressed } \sigma]}$: Don't delete a segment in the stressed syllable.
(b) MAX-Seg$_{[\text{initial } \sigma]}$: Don't delete a segment in the initial syllable.

The constraints in (65a–b) are not stipulated for the purposes of this analysis. They exist anyway because they are positional faithfulness constraints relativized to the two well-known LOCI of privilege (Trubetzkoy 1939): stressed syllables and initial syllables (see Beckman 1997, 1999; Casali 1997). Ranked above **ʔ*, MAX-Seg$_{[\text{stressed } \sigma]}$ and MAX-Seg$_{[\text{initial } \sigma]}$ thwart *ʔ*-deletion, as shown in (66). The example is *oce'an* [ʔɔ.tsɛ. ʔan] 'ocean', which exhibits glottal stops in both positions of privilege. The stressed syllable is indicated with an accent ˈ.

(66) *Level 3* /ʔɔ.tsɛ.ʔan / = [ʔɔ.tsɛ.ʔan]

| | MAX-Seg$_{[\text{stressed } \sigma]}$ | MAX-Seg$_{[\text{initial } \sigma]}$ | *ʔ | MAX-Seg | ONSET | ONSET-*w* |
|---|---|---|---|---|---|---|
| ☞ a. ʔɔ.tsɛ.ˈʔan | | | ** | | | |
| b. ʔɔ.tsɛ.ˈan | *! | | * | * | * | |
| c. ɔ.tsɛ.ˈʔan | | *! | * | * | * | |
| d. ɔ.tsɛ.ˈan | *! | * | | ** | ** | |

In addition to regulating the distribution of glottal stops, level 3 phonology spells out the glide /w/ as a labial approximant. The driver for the change is the segment inventory constraint **w*.

(67) **w*: Don't be [w].

This constraint plays no role at levels 1 and 2 because it is bottom-ranked. Importantly, **w* is ranked below IDENT[-cons], so consonantization cannot occur in the winning candidate.

(68) IDENT[-cons]
[-cons] on the input segment must be preserved on a correspondent of that segment in the output.

At level 3, the constraints are reranked to **w* >> IDENT[-cons], opening the way to /w/ changing into a different segment.



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

(69) Reranking at level 3
Level 2: IDENT[-cons] >> *w
Level 3: *w >> IDENT[-cons]

The mechanics of the *w* → *ʋ* spell-out need to be specified because /w/ should not surface as, for example, [z] or [r]. The spell-out is controlled by IDENT constraints, specifically, by IDENT-Lab, IDENT[+sonor] and IDENT[+contin].

(70) (a) IDENT-Lab: LABIAL on the input segment must be preserved on a correspondent of that segment in the output.
(b) IDENT [+sonor]: [+sonor] on the input segment must be preserved on a correspondent of that segment in the output.
(c) IDENT [+contin]: [+contin] on the input segment must be preserved on a correspondent of that segment in the output.

The ranking of these constraints is not relevant as long as they outrank IDENT [-cons]. For that matter, they could be the undominated constraints. The details of the evaluation for *ˈkanuw+a* 'canoe' GEN.SG., /ka.nu.wa/ → [ka.nu.ʋa], are laid out in (71).

(71) *Level 3* /ka.nu.wa/ → [ka.nu.ʋa]

| | IDENT-Lab | IDENT[+sonor] | IDENT [+contin] | *w | IDENT[-cons] |
|---|---|---|---|---|---|
| a. ka.nu.wa | | | | *! | |
| ☞ b. ka.nu.ʋa | | | | | * |
| c. ka.nu.va | | *! | | | * |
| d. ka.nu.ma | | | *! | | * |
| e. ka.nu.ja | *! | | | | |
| f. ka.nu.ra | *! | | | | * |

IDENT-Lab narrows down the pool of acceptable outputs to labials. IDENT [+contin] (or IDENT[-nas]) excludes candidate (71d) since nasals are [-contin]. IDENT[+sonor] makes sure that /w/ does not change into a labio-dental obstruent [v], which is what we see in (71c).

To conclude, level 3 phonology regulates the distribution of glottal stops and spells out /w/ as [ʋ].[54]

[54] A reviewer asks whether Harmonic Serialism (McCarthy 2000, 2007, 2010, 2016), a theory that is inherently derivational, can account for opacity and points out that even though McCarthy did



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

## 5. Conclusion

Upper Sorbian exhibits a complex Onset-driven conspiracy that involves a number of disparate processes: gliding into the onset, gliding into the coda, *j*-insertion, *ʔ*-insertion, and *w*-insertion. In some contexts, none of these processes can apply, so the optimal output is the one exhibiting hiatus. This is an impressively large number of surface configurations that the grammar is required to generate. The added difficulty is that, first, the processes in question operate differently in different constituents and, second, the derivation can be sensitive to properties that extend beyond segmental phonology by including crucial reference to the initial position and the occurrence in a stressed syllable. This conundrum is easily solved in Stratal/Derivational OT, which admits three levels of evaluation: the stem level, the word level and the post-lexical level. The arguments for levels are drawn not only from opacity, a classic source for such arguments, but also from cyclic effects and from the derivation of types of segments that are different at different levels.

The processing of words such as ˈ*rabij*+*a* [ija] GEN.SG. and ˈ*rabij*+*at*+*n*+*y* [ija] ADJ must be cyclic because it requires that syllable structure must be first assigned in the domain of the root before it is assigned in the domain of the word (see (54) in Section 4). The cyclic effect is derived because Upper Sorbian defines the stem at level 1 as the morphological root. This is new, not known about any other language, so Upper Sorbian is of interest from a typological perspective.[55]

Segment inventories are different at different levels of derivation. Specifically at levels 1 and 2, the epenthetic sonorant segments are /j/ and /w/. The /w/ turns into a consonant at level 3, a situation that is easily represented in Stratal/Derivational OT, but is impossible to represent in classic OT.

Directionality of glide insertion requires postulating a new constraint, *Multi, that assigns a violation to candidates containing the glide and its spawning vowel inside one syllable. This exactly is the reverse of what Itô & Mester (1999) meant to achieve with their Crisp Edge constraint.

Positional faithfulness is superior to positional markedness. The analysis has argued for Max-Seg relativized to stressed syllables and to word-initial syllables, exactly as predicted by positional faithfulness. Finally, counter to McCarthy (1999), Duke of York derivations are attested in OT and need to be recognized as legitimate, as is done by Stratal/Derivational OT.

## REFERENCES

Balas, Anna. 2011. Glottal stops produced by Polish native speakers in Polish and English. *Proceedings of the International Congress of Phonetic Sciences*, 280–283.

Beckman, Jill N. 1997. Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14, 1–46.

not envisage this function for Harmonic Serialism, the theory has been used exactly for this purpose by Torres-Tamarit (2012) and Hauser, Hughto & Somerday (2014). The matter awaits further investigation.

[55] As noted earlier, Spanish crucially defines the level 1 input not as a root but as a derived stem (Bermúdez-Otero 2013).



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press

Beckman, Jill N. 1999. *Positional faithfulness: An Optimality Theoretic treatment of phonological asymmetries*. New York and London: Garland Publishing.
Bermúdez-Otero, Ricardo. 1999. *Constraint interaction in language change: Quantity in English and Germanic*. Ph.D. dissertation, University of Manchester.
Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus* 25, 3–103.
Bermúdez-Otero, Ricardo. 2018. Stratal Phonology. In S. J. Hannah and Anna R. K. Bosch (eds.), *Routledge handbook of phonological theory*, 100–134. Abingdon: Routledge.
Casali, Roderick F. 1997. Vowel elision in hiatus contexts: Which vowel goes? *Language* 73, 493–533.
Charette, Monik. 1991. *Conditions on phonological government*. Cambridge: Cambridge University Press.
Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
Clements, George N. & Samuel J. Keyser. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
Halle, Morris. 1992. Phonological features. In William Bright (ed.), *International encyclopedia of linguistics*, 207–212. Oxford: Oxford University Press.
Hauser, Ivy, Coral Hughto & Megan Somerday. 2014. Faith UO: Counterfeeding in Harmonic Serialism. In A. Albright & M. A. Fullwood (eds.), *Proceedings of the Annual Meeting on Phonology 2014*, 2377–3324. Washington, DC: Linguistic Society of America.
Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
Itô, Junko & Armin Mester. 1999. Realignment. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody-morphology interface*, 188–217. Cambridge: Cambridge University Press.
Jocz, Lechosław. 2011. *Wokalowy system hornjoserbskeje rěče přitomnosće*. Szczecin: Volumina.
Kawasaki, Haruko. 1982. *An acoustical basis for universal constraints on sound sequences*. Ph.D. dissertation, University of California at Berkeley.
Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2, 85–138.
Kiparsky, Paul. 1997. *LP and OT. Handout at LSA Summer Linguistic Institute*, Cornell University.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17, 351–365.
Levin, Juliette. 1985. *A metrical theory of syllabicity*. Ph.D. dissertation, MIT.
Maćijowa, Maria. 2007. Nowe postajenja na polu ortografije, morfologije a interpunkcije w hornjoserbšćine. *Unpublished ms*., Domowina Verlag, Bautzen.
McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16, 331–399.
McCarthy, John J. 2000. Harmonic Serialism and parallelism. *Proceedings of the North East Linguistics Society* (NELS) 30, 501–524.
McCarthy, John J. 2007. *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox Publishing.
McCarthy, John J. 2010. An introduction to Harmonic Serialism. *Language and Linguistics Compass* 4, 1001–1018.
McCarthy, John J. 2016. The theory and practice of Harmonic Serialism. In John J. McCarthy & Joe Pater (eds.), *Harmonic grammar and Harmonic Serialism*, 47–87. London: Equinox Publishing.
McCarthy, John J. & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18, 249–384. Amherst, MA: GLSA Publications.
Michalk, Frido. 1955. *Gramatika. Serbšćina: Listowy studij za wučerjow*. Bautzen: Domowina Verlag.
Prince, Alan & Paul Smolensky. [1993] 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell. [Revision of 1993 technical report, Rutgers University Center for Cognitive Sciences. Available on Rutgers Optimality Archive, ROA-537.]
Pullum, Geoffrey. 1976. The Duke of York gambit. *Journal of Linguistics* 12, 83–102.
Rubach, Jerzy. 1997. Extrasyllabic consonants in Polish: Derivational Optimality Theory. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 551–581. Oxford: Oxford University Press.
Rubach, Jerzy. 2000a. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31, 271–317.
Rubach, Jerzy. 2000b. Backness switch in Russian. *Phonology* 17, 39–64.
Rubach, Jerzy. 2002. Against subsegmental glides. *Linguistic Inquiry* 34, 672–687.
Rubach, Jerzy. 2003. Duke of York derivations in Polish. *Linguistic Inquiry* 34, 601–629.
Rubach, Jerzy. 2008. Palatal nasal decomposition in Slovene, Upper Sorbian and Polish. *Journal of Linguistics* 44, 169–204.
Rubach, Jerzy. 2011. Syllabic repairs in Macedonian. *Lingua* 121, 237–268.

Rubach, Jerzy. 2014. Soft labial conspiracy in Kurpian. *Journal of Linguistics* 50, 185–230.
Rubach, Jerzy. 2016. Polish yers: Representation and analysis. *Journal of Linguistics* 52, 421–466.
Rubach, Jerzy. 2019a. Three arguments for underspecified representations. *Studies in Polish Linguistics* 14, 191–217.
Rubach, Jerzy. 2019b. Surface Velar Palatalization in Polish. *Natural Language and Linguistic Theory* 37, 1421–1462.
Sagey, Elizabeth. 1986. *The representation of features and relations in non-linear phonology*. Ph.D. dissertation, MIT.
Schaarschmidt, Gunter. 2002. *Upper Sorbian*. Munich: Lincom Europa.
Schuster-Šewc, Hinc. 1968. *Gramatika hornjoserbskeje rěče*. Bautzen: Domowina Verlag. Translated into English by Gary H. Toops as *Grammar of the Upper Sorbian Language* (1996). Munich: Lincom Europa.
Schuster-Šewc, Hinc. 1983. *Historisch-etymologisches Wörterbuch der ober- und nieder-sorbischen Sprache*. Bautzen: Domowina Verlag.
Schwartz, Geoffrey. 2012. Initial glottalization and final devoicing in Polish English. *Research in Language* 10, 159–171.
Stone, Gerald. 1993. Sorbian (Upper and Lower). In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 593–685. London: Routledge.
Torres-Tamarit, Francesc. 2012. *Syllabification and opacity in Harmonic Serialism*. Ph.D. dissertation, Universitat Autónoma de Barcelona.
Trubetzkoy, Nikolai. 1939. *Grundzüge der Phonologie*. Translated into English as *Principles of phonology* (1969). Berkeley: University of California at Berkeley.
Van Oostendorp, Marc. 2007. Derived Environment effects and consistency of exponence. In Blaho, Sylvia, Patrick Bye & Martin Krämer (eds.), *Freedom of analysis?*, 123–148. Berlin: Mouton de Gruyter.
Völkel, Pawoł & Timo Meškank. 2005. *Prawopisny słownik hornjoserbskeje rěče*. Bautzen: Domowina Verlag.
Wowčerk, Pawoł. 1955. *Kurzgefasste obersorbische Grammatik*. Berlin: Volk und Wissen Ludowy Nakład.
Żygis, Marzena, Janna Brunner & Scott Moisik. 2012. The affinity of low vowels and glottal stops: An articulatory model and a perceptual investigation. *Proceedings of the 13th Conference on Laboratory Phonology*, 27–28.

*Author's address:* *Department of Linguistics,*
*University of Iowa, Iowa City, IA 52242, USA*

*Department of Linguistics, University of Warsaw.*
*jerzy-rubach@uiowa.edu*



https://doi.org/10.1017/S0022226722000421 Published online by Cambridge University Press