An exploration of sonorant-obstruent straddlers

Mishko Bozhinoski
2023

UNIVERSITY OF AMSTERDAM

1. Introduction

Most modern phonological literature bases itself on the notion that contrastive sound units can be described as a constellation of discrete acoustic or articulatory characteristics. The sound [p], for instance, can be described as being “labial”, “plosive”, “voiceless”, etc. Sounds which share certain characteristics are expected to be subject to similar distributional or transformational rules, and can thereby be grouped in classes or categories. As the sounds [m], [n] and [ŋ] are all nasal stops, one expects then that they act similarly in a given language. Within most feature-based theories, the properties (features) that determine the category of a sound are taken to be binary. The feature [+-obstruent], thus, represents partial or complete turbulent occlusion, with [+obstruent] representing the presence of occlusion (thus, an obstruent) and [-obstruent] representing the lack thereof (thus, a sonorant). This entails trivially that a sound cannot be both an obstruent and not a sonorant at the same time.

Yet, speech sounds characterizing both sonorants and obstruents are frequently observed. Sounds that involve fricative-like occlusion are often observed to pattern unambiguously with sonorants, such as German [v] and French [ʁ] (see section §2.2). Furthermore, one finds sounds that pattern with both sonorant and obstruents at the same time: Russian [v] devoices before voiceless obstruents and word boundaries (like voiced obstruents) yet does not cause preceding voiceless obstruents to voice (like sonorants). While this sound has brought on much attention from phonologists, I suspect that it is far from unique in the languages of the world. In §2.3, I give a similar example in Belgian Dutch. The topic of the present undertaking concerns this latter category of *straddlers* – consonants that simultaneously behave as sonorants and obstruents and thus straddle the boundary between them.

In a typological study on the phonological status of voiced fricatives, Botma & van ‘t Veer (2013) state that many voiced non-sibilant fricatives with no voiceless counterpart are in fact best categorized as sonorants (the reasoning for which I outline in §2.3). The present research expands on this by putting forward the following hypothesis: many voiced non-sibilant fricatives with no voiceless counterpart are in fact straddlers, showing independent class behavior with both sonorants and obstruents. In light of this hypothesis, the specific aim of this study is twofold: firstly, to test this hypothesis by examining the specific class behavior of these obstruents through a typological corpus study; and secondly, if any straddlers are found, to document in which way(s) they pattern with sonorants and obstruents, with as a sub-goal to look for any patterns in the patterns.

Below are the assumptions upon which this research is based:

Firstly, as outlined at the start of this section, sounds form categories on the basis of their shared phonological behavior, such as being subject to distributional constraints, transformational rules or being part of a conditioning environment. This assumption is outlined in more detail in §2.1.

Secondly, sounds that tend to pattern together across different languages are expected to do so in similar ways. For instance, [p t k] are expected to act similarly in both language A and language B provided both languages A and B have the category [p t k].

Finally, sounds that form a behavioral class are assumed to also share some phonetic substance. In other words, sounds that can be grouped based on behavior are expected to have a common acoustic or articulatory thread. Similarly, sounds with a common phonetic property are also expected to behave similarly. Substance-free approaches are acknowledged, but shown to face much of the same issues posed by the straddler problem. This is outlined in §2.4.

2. Background

2.1 Phonological categories and the sonorant-obstruent dichotomy

Within contemporary phonology, it is taken for granted that phonological rules operate on categories rather than individual sounds. To demonstrate the inadequacy of proposing rules that operate on individual sounds, take the following example. In Dutch, the consonants /d b z v ɣ/ are rendered [t p s f x] respectively in word-final position (Booij, 1999, p. 60), as exemplified in (1).

(1)	**Infinitive**	**1sg**	**Meaning**
	[ɦɛ**b**ən]	[ɦɛ**p**]	‘have’
	[rɛi**d**ən]	[rɛi**t**]	‘ride’
	[ʋɛi**z**ən]	[ʋɛi**s**]	‘show’
	[ɣeː**v**ən]	[ɣeː**f**]	‘give’
	[zɛ**ɣ**ən]	[zɛ**x**]	‘say’

We may thus formalize this rule as such:

/d b z v ɣ/ > [t p s f x] / _#

This misses a few key insights, however. Firstly, it misses the fact that all sounds affected are voiced obstruents; and secondly, it misses that the resulting consonants are all of the same place and manner of articulation as their pre-rule form, with the only difference being that they are all voiceless. It is preferable then to phrase this in terms of categories: “voiced obstruents are devoiced word-finally”. This is not only more concise, but also more insightful. It predicts, for instance, that if Dutch were to acquire a new voiced obstruent, it would devoice it at the end of a word. This is exactly what we see in recent loanwords, such as the English loan [dʒɔgən] ‘to jog’, in which the acquired non-native [g] devoices to [k] in forms such as [ɪk dʒɔk] ‘I jog’. Through this, we establish two things. Firstly, we establish that the Dutch rule outlined above applies not to an arbitrary array of sounds, but to one specific category described in terms of abstracted properties (namely voicing and obstruency). Secondly and as a result, we establish that voiced obstruents is a relevant category in Dutch, and that /d b z v ɣ/ are sounds that comprise it. We observe from such examples that human language strives for computational efficiency through representational symmetry: one would not expect a hypothetical plosive /g/ *not* to devoice to [k], as that would not be in line with other sounds of its category, requiring a more complex rule governing the distribution of obstruents. These were the ideas mainstreamed by the seminal *The Sound Pattern of English (Chomsky & Halle, 1968)*, which has pushed the notions of feature-based categorization and its relationship to computational efficiency.

Another way in which one may establish the category of a sound or group of sounds is through its distribution. In almost all (if not all) languages, there is a limitation to which sounds are allowed to occur in a given skeletal position. It is unexpected that a given language will allow all sounds to occur in all positions offered by the language within a word, syllable, foot, root, etc. A common such constraint concerns the structure of consonant clusters. In a great many languages around the world, boundary clusters follow a very rigid structure, namely that the consonant closer to the vowel must involve less oral occlusion than the consonant farther away from it (Clements, 1990). In other words, the closer a member of a cluster is to the vowel, the more ‘vowel-like’ it must be.

(2)	**Thai**	[**kl**ɔ̄ːŋ]	‘drum’	*[**lk**ɔ̄ːŋ]
	Mauritian Creole	[**tr**ɛz]	‘thirteen’	*[**rt**ɛz]
	Turkish	[re**nc**]	‘color’	*[re**cn**]
	Dutch	[dɑ**mp**]	‘vapor’	*[dɑ**pm**]

By this distribution, we observe that the sounds [p k t c] (among others) occupy positions farther from the vowel, and that the sounds [l r n m] (among others) occupy positions closer to the vowel. This is one of the bases for establishing the sonorant-obstruent distinction, a distinction fundamental to segmental phonology. The former sounds, thus the sounds that involve more occlusion, are referred to as *obstruents*, while the latter are referred to as *sonorants*.

2.2 Phonetic challenges to the sonorant-obstruent dichotomy

This distinction between sonorants and obstruents on the basis of acoustic and/or articulatory properties presents many problems once we attempt to square it with what is actually observed in languages. A phonetic property usually used to identify obstruents is turbulent occlusion – the complete or partial obstruction of air within the oral cavity in such a way that it generates aperiodic high-frequency noise. Yet, it is not uncommon for speech sounds that involve turbulent oral occlusion to function as sonorants in a given language. A common such case are many rhotics, which may involve highly turbulent constriction despite having a sonorant-like distribution. A well-known case of this is French, whose rhotic is typically realized as a highly turbulent uvular fricative [ʁ] or [χ] yet occupies sonorant positions within consonant clusters.

(3)	[kle] ‘key’	*[lke]
	[pli] ‘fold’	*[lpi]
	[kχi] ‘cry’	*[χki]
		*[ʁli]

The same speech sound may in other languages be interpreted as a fricative. In Northern Dutch, for instance, [χ] patterns unambiguously with fricatives.

(4)	[klɛi] ‘clay’	*[lkɛi]
	[χlɛi] ‘glide’	*[lχɛi]
	[bɑlk] ‘beam, plank’	*[bɑkl]
	[bɑlχ] ‘bellows’	*[bɑχl]

A similar pattern is observed in German [v], which involves fricative-like oral occlusion yet still occupies sonorant positions ([k^hvaːl] ‘agony’ *[vk^haːl]). Just as with the French rhotic, [v] in Dutch represents an ambiguous fricative, as shown in words such as [vrɑu] ‘woman’.

The French and German examples presented above show that turbulent occlusion, the primary phonetic property associated with obstruents, is unreliable in establishing the sonorance of a given sound as a phonological category. This conclusion is further borne out in the typological literature on the phonology of sonorants.

2.3 Typological challenges to the sonorant-obstruent dichotomy

In 2013, Botma & van ‘t Veer conduct a typological study on voiced non-sibilant fricatives that sheds light on this issue. They note that voiced fricatives do not follow the typological distribution expected of true obstruents, whose unmarked state is to be voiceless. In other words, a language that has voiced plosives or sibilant fricatives will also have voiceless plosives. Thus, a language that has [d g z] is expected to also have [t k s]. Voiced non-sibilant fricatives do not follow this pattern, and in fact follow the exact opposite pattern: they occur *more* frequently without their voiceless counterpart than with, a pattern more familiar to sonorants. On this basis, they hypothesize that many of these offending fricatives are best analyzed as sonorants.

To investigate this hypothesis, they conduct a study on 70 languages with unpaired voiced fricatives through examining their class behavior as documented in P-Base (Mielke, 2008), a typological database that documents the phonological rules and distributions of 537 languages. Out of the 70 languages they document, they find that 13 of them have their unpaired voiced fricatives pattern exclusively with sonorants. In 15 languages, they observe behavior that targets a class containing both fricatives and sonorants, making the grouping of the unpaired fricatives unclear. In 34 of the 70 languages, phonological activity was specific to the unpaired voiced fricatives, and thus inconclusive. Finally, the last 8 languages showed patterning with obstruents, which runs counter to their hypothesis. Upon closer examination, these prove to either support their classification as sonorants[1] or to provide no conclusive counter-evidence to the hypothesis that unpaired voiced fricatives are best analyzed as sonorants.

As introduced in §1, I offer a modification to this hypothesis: at least some of these offending fricatives show class behavior of *both* sonorants *and* obstruents at the same time. In other words, they show phonological activity that places them in the sonorant category while at the same time showing different phonological activity that places them in the obstruent category. The next section outlines two examples of this.

2.4 Straddlers – a phonological challenge to the sonorant-obstruent dichotomy

So far, the discussion challenges the sonorant-obstruent distinction on the basis of the discrepancy between their phonetic properties versus their phonological patterning. The assumption has been, thus, that sounds that behave alike have some shared phonetic substance (i.e. acoustic or articulatory properties). This assumption is not universal, and is rejected by proponents of the aptly-named *substance-free phonology*, whose adherents propose that phonology does not contribute nor operate on phonetic properties – phonological categories are thus not to be based on phonetic implementation (see Mielke, 2005; Blaho, 2008; Reiss, 2017). While on the surface this seems to provide a solution to the problem, below are cases that pose a problem even if one were to completely disregard phonetic substance.

2.4.1 Russian /v/

Russian /v/ has garnered a lot of attention from phonologists due to its peculiar voice assimilation patterns (Andersen, 1969; Coats & Harshenin, 1971; Iosad, 2018; Lulich, 2004; Padgett, 2002; Petrova & Szentgyörgyi, 2004). Firstly, /v/ is rendered as [f] word-finally, just like all voiced obstruents and unlike sonorants.

[1] By, for instance, showing that they are the result of intervocalic lenition, which they argue is best viewed as sonorization

(5)

	Genitive sg.	Nominative sg.	Translation
	sleda	slet	track
	raza	ras	occasion
	plʲaʒa	plʲaʃ	beach
	lʲubvʲi	lʲubofʲ	love
	vola	vol (*vol̥)	ox

Secondly, /v/ assimilates in voicing to any following obstruent, just like all obstruents.

(6)

Preposition	Before voiced	Before voiceless
iz ‘out of’	is papɨ ‘out of dad’	iz babuʃki ‘out of grandma’
k ‘to’	k tʃeloveku ‘to the man’	g dime ‘to Dima’
v ‘in’	f tʃeloveku ‘in the man’	v dima ‘in Dima’

From this data alone, one may conclude that /v/ is an obstruent, as it is targeted by the same rule that targets other obstruents such as /k g z d/ and so on. However, looking at the voice assimilation again, but this time as the role of assimilator rather than assimilatee, we see that /v/ does not cause previous voiceless obstruents to voice (nor devoice), a property only seen in sonorants.

(7)

Preposition	Before /v/	Before sonorant
ot ‘from’	ot vas (*od v-) ‘from you’	ot mzdɨ ‘from the bribe’
pod ‘under’	pod vas (*pot v-) ‘under you’	podnesti ‘to bring to’
s ‘with’	s vami (*z v-) ‘with you’	s mamoj ‘with mama’

Additionally, looking at word-initial C_1C_2 clusters, we see that Russian permits [v] to occupy C_2 even when C_1 is a voiceless obstruent, a pattern that is otherwise only found in sonorants and other voiceless obstruents.

(8)

Word	Meaning
plavatʲ	‘to swim’
krof	‘shelter, roof’
sverx	‘above’
skvaʒina	‘chink’

From the data of (7) and (8), then, one may instead conclude that /v/ is a sonorant. Thus, it is perfectly possible to both claim that /v/ is a sonorant and claim that it is an obstruent, two categories traditionally viewed as polar opposites, and therefore mutually exclusive.

This pattern in /v/ is not only found in Russian, but other Slavic languages. In Macedonian, we also see the existence of words such as [svoj] ‘one’s own’ categorize /v/ as a sonorant, while we also observe /v/ to assimilate in voicing with following obstruents, as in /v-klutʃi/ [fklutʃi] ‘turn on’ (Friedman, 2002, p. 12).

2.4.2 Belgian Dutch /ɣ/

In Standard Dutch, and thus also Belgian Dutch, voiced obstruents devoice in word-final position (see §2.1). They additionally assimilate in voicing with following plosives.

(9)	**Word**	**Before sonorant**	**Before obstruent**
	[ɪk] 'I'	[ɪk leːf] 'I live'	[ɪg bɛn] 'I am'
	[ɔp] on	[ɔp maːt] 'bespoke'	[ɔb dɪt momɛnt] 'at this moment'
	[lees] 'read!'	[leːs niet] 'do not read'	[leːz də tɛkst] 'read the text'

The voiced dorsal continuant /ɣ/ joins the obstruents in undergoing final devoicing and voice assimilation.

(10)	**Infinitive**	**1sg present**	**1sg preterite**	**Meaning**
	[zaːɣən]	[zaːx]	[zaːɣdə]	'complain'
	[vɔlɣən]	[vɔlx]	[vɔlɣdə]	'follow'
	[ɛːndəɣən]	[ɛːndəx]	[ɛːndəɣdə]	'finish'

Furthermore, /ɣ/ patterns as an obstruent in being able to occupy the first position of a complex onset.

(11)		
	[droːm]	'dream'
	[trɑp]	'step, stairs'
	[vraːx]	'question'
	[ɣroːt]	'big'

Finally, /ɣ/ patterns with voiced fricatives in particular in being subject to devoicing after voiceless obstruents (Booij, 1999, p. 58).

(12)		
	[zɪŋt] 'sings'	[ət sɪŋt] 'it sings'
	[vɪnt] 'finds'	[ət fɪnt] 'it finds'
	[ɣaːt] 'goes'	[ət xaːt] 'it goes'

From this data set, we may conclude that /ɣ/ is a voiced obstruent, more precisely a voiced fricative.

Looking at the word-medial distribution of /ɣ/, however, reveals a peculiarity. Belgian Dutch, like Standard Dutch, has a tight co-occurrence pattern with regard to the vowel-fricative sequences (van Oostendorp, 2003, p. 330). Intervocalically, voiceless fricatives may only occur after short vowels, and voiced fricatives may only occur after long vowels.

(13)		
	[ɔntplɔfən] 'to explode'	*[ɔntplɔvən]
	[lɛsən] 'lessons'	*[lɛzən]
	[lɑxən] 'to laugh'	*[laːxən]
	[ɣeːvən] 'to give'	*[ɣɛvən]
	[leːzən] 'to read'	*[lɛzən]

Yet, /ɣ/ can be seen to occur both after short and long vowels.

(14)		
	[zeːɣən] 'blessing'	[zɛɣən] 'to say'
	[vlɑɣən] 'flags'	[vlaːɣən] 'gusts'

The voiced velar continuant, thus, shares this distribution with sonorants and plosives. Given that the voice assimilation pattern of /ɣ/ does not follow that of stops, and that we hear word-final /ɣ/ voiced across word boundaries (/zɛɣ ɪk/ [zɛɣ ɪk] 'I say' vs /raːd ɪk/ [raːt ɪk] 'I guess'), we must conclude that intervocalic /ɣ/ patterns with sonorants.

Russian /v/ and Flemish Dutch /ɣ/ present a challenge not just to the mapping between the phonetic concept of occlusion and the more abstract phonological categorization of sonorance, but to the phonological categorization itself. They show direct patterning with both the class of sonorants and that of obstruents separately, and thus straddle the boundary between them. Henceforth, these sounds will be referred to as *straddlers*. This puts in question the current conceptualization of class membership of sonorants and obstruents, which presupposes that belonging to the class of sonorants necessarily means *not* belonging to the class of obstruents. Straddlers push us to either dispense with this idea, either by reconceptualizing sonorance-based class membership or dispensing of sonority altogether.

3. Method

To reiterate on the hypotheses offered at the end of §1 and §2.3, I suspect that many voiced obstruents with no voiceless counterpart, whose distribution runs counter to typological expectations, are in fact straddlers. The aim, then, is to test this hypothesis through a typological corpus study, and, if straddlers are found, to document any patterns found within their behavior

The methodology of this research consisted of three stages: language selection, pre-analysis preparation and analysis. Each of these three stages is explained below. All code utilized for this project is available on *github.com/jassummisko/Straddlers*.

3.1 Selection phase

The first phase involved the selection of the pool of languages to be investigated for potential straddlers. The database used was Phoible (Moran & McCloy, 2019), which compiles more than 3,000 phonemic inventories and categorizes their sounds in terms of distinctive features, such as [+-nasal] or [+-high]. It additionally lists (when the primary source for the language provides it) a list of allophones of each phoneme.

First, the selection criteria for the selection of fricatives were devised (represented as a flowchart in Figure 1). Based on the heuristics utilized in Botma & van't Veer (2013), the first criterion was whether the language had any voiced obstruents that do not have a voiceless pair. Voiced obstruents whose voiceless equivalents were marked as being marginal were also to be included in the dataset. However, languages whose gaps are typologically commonplace (such as the presence of /b/ but no /p/, as in most Arabic varieties) were not included in the dataset.

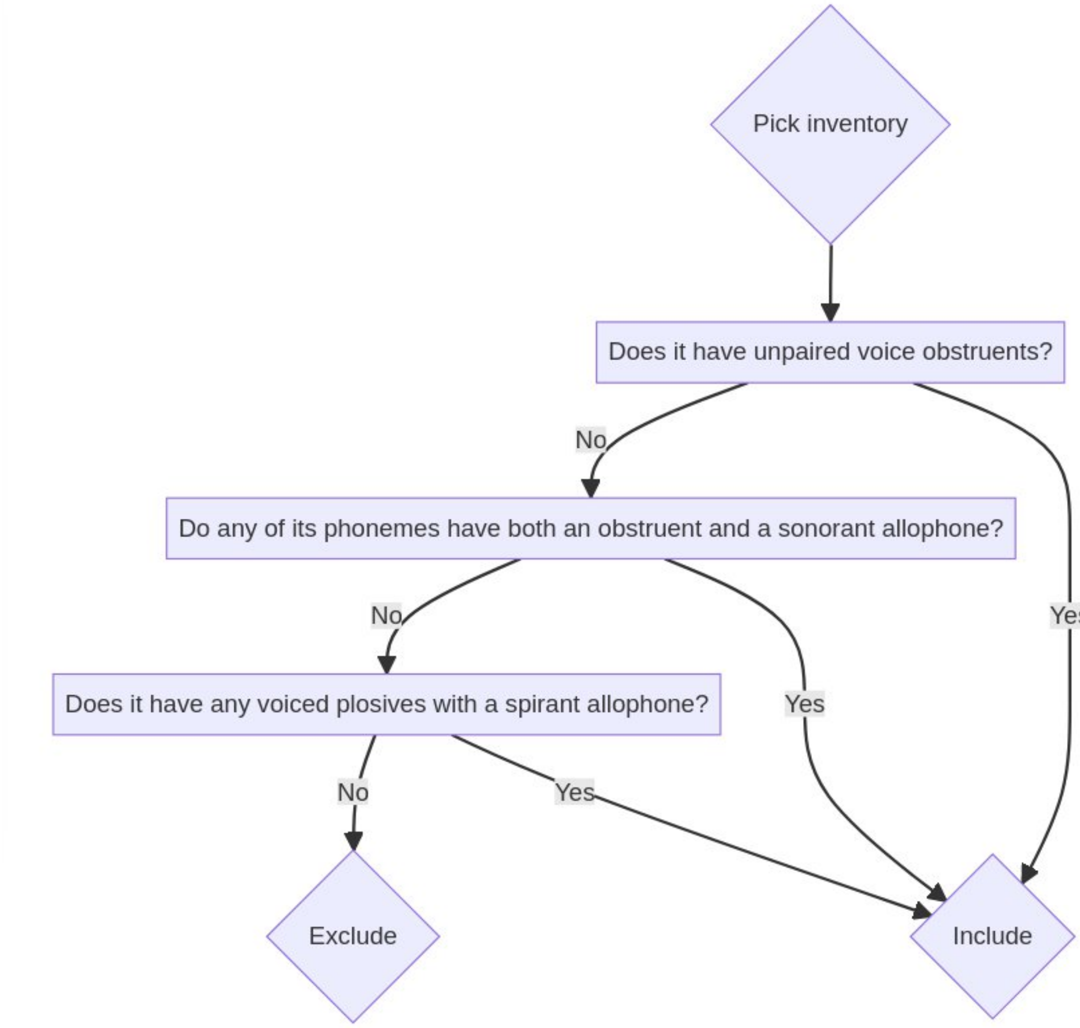
The second criterion was whether any of the languages' phonemes had both a sonorant and an obstruent allophone. For instance, if a language has the phoneme /b/ with the allophones [b w], then it was included in the pool.

Finally, the third criterion was whether any of the voiced plosives of the language in question had a spirantized allophone. Thus, if a language had the phoneme /d/ with the allophones [d ð], then it was also included in the pool.

Once the selection criteria were established, a program was written in *Yuescript* (*github.com/pigypigyyy/Yuescript;* last accessed May 2023) that processes the feature and inventory data provided by Phoible. A brief summary of the workings of the program is as follows. First, all segment characters used by Phoible are stored together with their corresponding feature values as a hash table with keys corresponding to the name of the feature. Then, all inventories are read with together their corresponding segments. Finally, each segment of each inventory is checked against the aforementioned three criteria.

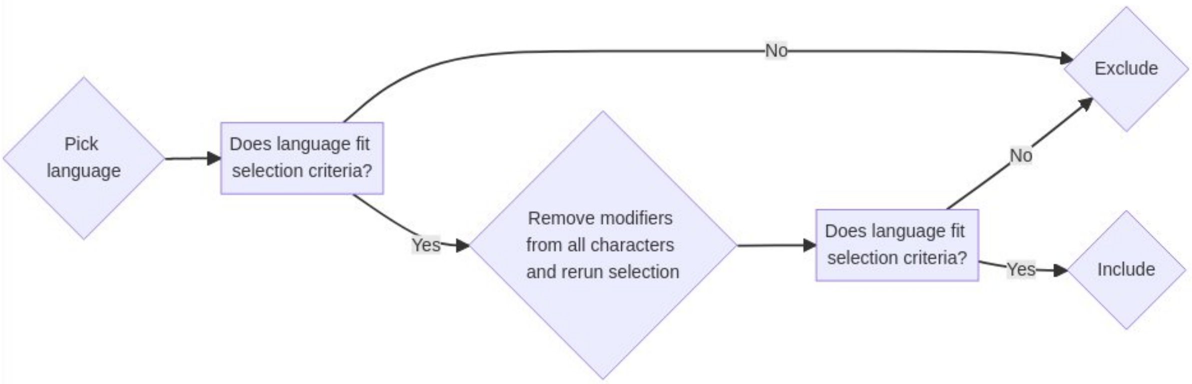
The initial algorithm resulted in many false-positives due to the problematic nature of how Phoible represents its segments. Namely, it is assumed that segments with the same phonetic value across multiple languages also have the same featural representation. To give an example, a [pʰ] in one language is treated as being the same as [pʰ] in another. This is problematic in the context of this research, as [pʰ] may fill multiple different functions from language to language. In some language A with a two-way VOT distinction in plosives, [pʰ] may be the voiceless equivalent to [b]. In language B, in which there is a three-way distinction, [pʰ] may function as an aspirate in the labial plosive series [pʰ p b]. In language C, where there is no VOT distinction in obstruents, [pʰ] may simply be the only labial plosive that happens to be accompanied by aspiration. In language A, the feature [+-voiced] is sufficient to distinguish [pʰ] from [b]. In language B, the feature [+-spread glottis] is necessary alongside [+-voiced] to make all necessary VOT distinctions in plosives. In language C, neither [+-spread glottis] nor [+-voiced] are necessary and may simply be unspecified.

Figure 1: Inventory selection process



This problem was mitigated by augmenting the original selection algorithm with a unicode normalization step, presented graphically in Figure 2. Briefly: the algorithm first checks if an inventory falls under the criteria outlined above. If it does, the characters used for each of its phonemes are normalized in such a way that all of its modifiers and diacritics are not taken into account (thus, the characters <*pʰ*>, <*p'*>, <*p*>, and so forth are able to be counted as being the same). If the language still fits the selection criteria after normalization, it is included in the final set.

Figure 2: Normalization step



This ameliorates the problem of false-positives by solving languages that contrast, say, [pʰ] vs. [b] or [s] vs. [z̤]. This, however, only mitigates the problem and does not completely solve it; false positives were still spotted.

The final algorithm resulted in 1014 inventories selected in total, which is around a third of all inventories stored on Phoible. As there were strict time constraints on this research project and the method was to involve manual checking of the sources for each language, the algorithm was rerun to include only voiced fricatives, which lowered the number to 244 inventories. Finally, each language was manually checked and excluded if it was a false positive, i.e. did not actually meet any of the three selection criteria. This lowered the language pool to around 200 inventories, which is the pool that was used when proceeding to the next phase.

3.2 Scraping and source collection phase

After the manual check, the primary sources for each language were retrieved when possible. During this phase, many languages were excluded for a multitude of reasons (see appendix for excluded languages with a reason why), which are outlined below:

1. The primary source for the language was not able to be retrieved.

2. The primary source did not contain any information necessary for the paper, such as phonological patterning or information on morphophonological alternations. Most commonly, these were dictionaries with a short preface containing the phonological inventory of the language and the orthographical representation of each segment.

3. The primary source was in a language that was not in a language understood by the author of this paper, thus not English, Dutch or a South Slavic language.

Finally, areal and genetic information on each language was scraped from Phoible using *Python* and the *requests* library (source available on *github.com/jassummisko/Straddlers*), and then the final analysis phase was commenced.

3.3 Analysis phase

The analysis phase consisted of retrieving relevant information from each source and collecting it in a database. This database consisted of one entry per segment, populated by the segment itself, the language it comes from, ten binary properties, and an optional field for extra notes where relevant. The binary properties had the possible values + (for 'yes'), - (for 'no'), and *0* (when not applicable, elaborated below). An overview of the binary selection criteria are represented in Table 1.

Languages whose source did not contain enough information relevant to these criteria were excluded (see appendix for all languages excluded with reasoning why). Additionally, as the goal of this research is to examine sonorant-obstruent straddlers, segments that were encountered during the analysis phase that showed straddling properties were included even if they were not detected by the algorithm outlined in §3.2.

Table 1: binary selection criteria

Does the segment...	**Not applicable if ...**
... appear word-initially?	... word-initial consonants are not allowed[1]
... appear as word-medial onset?	... word-medial onsets are not allowed[2]
... occupy the first slot of a complex onset?	... no complex onsets permitted
... occupy the second slot of a complex onset?	... no complex onsets permitted
... appear word-finally?	... no word-final consonants permitted
... appear as a word-medial coda?	... codas are not permitted
... occupy the first slot of a complex coda or word-final cluster?	... neither clusters permitted
... occupy the second slot of a complex coda or word-final cluster?	... neither clusters permitted
... pattern with obstruents? (including note how)	... not enough information in source
... pattern with sonorants?	... not enough information in source

4. Results

The initial data-set consisted of 122 segments spanning 83 languages. Before analysis, the data was manually filtered for false-positives, false-negatives and other reasons for inclusion or exclusion. Below is an outline of all of those reasons (see appendix for all languages excluded):

Reason	**Explanation**
False-positive	The language does have a voiceless equivalent of the segment in question that was not listed on Phoible.
False-negative	The language contains an unpaired voiced fricative that was not captured by the algorithmic filter.
Poor source	The source in question contains insufficient information.
Phoible error	The language does not have a segment listed as present in Phoible.

After the manual filtering process, the final data-set consisted of 102 segments spanning 58 languages.

17 segments patterned as obstruents only, spanning 15 languages and 9 unique segments.
35 segments patterned as sonorants only, spanning 26 languages and 12 unique segments.
23 segments patterned with both, spanning 11 languages and 10 unique segments.
2 segments of 2 languages showed inconclusive patterning
25 segments could not be placed in any class due to a lack of relevant information in the source

[1] The Australian language Oykangand allegedly does not allow onsets, including word-initially.

4.1 Patterns with obstruents

The following section shows all segments that were found to pattern exclusively with obstruents. See Table 2 for a representation of all such segments found with a per-language distribution of each.

Table 2: distribution across languages of obstruent-patterning segments in dataset

Segment	n	Found in:
ʁ	2	Muna, Sakha
v	1	Mavea
ɣ	3	Burushaski, Korafe, Kensei Nsei
β	4	Tigak, Ndebele, Emmi, Oykangand
ð	3	Mixtec, Rigwe, Cwaya
dʒ	1	Korafe
z	1	Kurmuk (Burun)
ʐ	1	Emmi
w	1	Muna

Below is an outline of the patterns observed within the set of consonants that show class behavior with obstruents.

4.1.1 Pairs with stop

Among the offending segments that only pattern with obstruents, almost half pair, pattern or alternate with stops.

In Sakha (Tolkov, 2011), Burushaski (Morgenstierne, 1945, p. 24) and Mocovi (Grondona, 1998, p. 100) all share in common that they have a dorsal fricative that pairs with a voiceless dorsal stop in devoicing contexts. In Mocovi, for instance, [ʁ] alternates with [q] in word-final position (where final obstruent devoicing applies), suggesting that this is the voiced equivalent of [q].

(15)	/s-amaʁ/	[sama**q**]	1sg-‘push’
	/amaʁ-i/	[ma**ʁ**ai]	‘push’-2sg.f

In Ndebele (Sibanda, 2004, p. 163), [β] seems to in fact pattern with ejectives, in that it functions as glottalic within root well-formedness rules. Ndebele root-initial C_1VC_2- may only contain an ejective in C_2 if C_1 is also an ejective. The only phonetic non-ejective that breaks this rule is [β].

(16)	**t’**o**t’**-	‘sink down in wet place’
	p’a**t’**ul-	‘slap lightly’
	βo**t’**oɣ-	‘be soft when felt with fingers’
	*po**t’**oɣ-	illegal root

Sibanda also claims that one sees a similar patterning with [ɣ], where it is in complementary distribution with [k’], but he gives no direct examples of this. It is worth noting that modern [β] in Ndebele corresponds to an implosive in closely related languages.

In Oykangand (Sommer, 1969, p. 35), [β] patterns with stops in VC_1C_2 syllables where C_1 is a nasal, being the only non-stop allowed to occupy the C_2 position. This could suggest it being a voiced equivalent of [p] if it were not for the fact that it actually contrasts with both [b] and [p] in this position.

4.1.2 Variation or patterning with voicelessness

In both Mavea (Guerin & Aoyama, 2009, p. 253) and Emmi (Ford, 1998, p. 42), /v/ and /β/ (respectively) has the usual realizations [v] and [β], but they may also be voiceless. In the case of Emmi, it is voiceless in word-initial positions. In the case of Mavea, voicing is always optional, rendering [v] and [f] free variants. In this case, Botma & van ‘t Veer (2013) suggest that these fricatives may not be specified for voicing at all, and are simply “lenis”.

A language that presents patterning with voicelessness despite phonetic voicing is Muna. This is discussed in detail in the next sub-section.

4.1.3 Case study: Muna [ʁ] and [w]

Roots in Muna (van den Berg, 1989, p. 30) have the constraint that homorganic obstruents within the same root must share the same voicing. Thus, that a word that contains [t] may only contain an alveolar obstruent if it is also voiceless, in this case, [t] or [s] (note that, as this constraint only applies to obstruents, there is also nothing stopping [t] and [l] from co-occurring within a root). This reveals two unusual patterns.

Firstly, the dorsal fricative [ʁ] only co-occurs with voiceless dorsals despite being phonetically voiced. Roots of the form [kVʁV] are therefore allowed, while roots of the form [ɡVʁV] are not. This seems to suggest that at some level of representation, [ʁ] is voiceless. It is worth noting that Muna [ʁ] is in fact a reflex of a voiceless uvular plosive */q/ in Proto-Austronesian, which provides a diachronic explanation of the current plosive distribution in roots.

Secondly, as the aforementioned root-internal voicing constraint applies only to obstruents, one would expect [w] to not participate in the constraint. Yet, it does: [w] may not co-occur with voiceless labial obstruents. This suggests that [w] is an obstruent despite its approximant realization.

As we are dealing with a root-level constraint, there is a possibility that the lack of co-occurence of [ʁ] and [w] with voiced dorsals and voiceless labials respectively is a statistical accident, or that it is a reflection of a no longer active historical rule that enforced like-voicing in roots at a point when these consonants had a different phonological status. More synchronically active evidence comes from their morphophonological alternations triggered by a number of prefixes that cause fortition and subsequent prenasalization of the initial consonant of the root, a process that van den Berg (1989, p. 35) dubs ‘nasal accretion’. The derivational morpheme ka- -ha is an example of such prefix, denoting a tool used to perform the action of the verb it attaches to:

(17)	[**t**ei]	‘to put, store’	[ka(e)-**nt**ei-ha]	‘vessel, container’
	[**f**umaa]	‘to eat’	[ka-**mp**oroʁu-ha]	‘eating utensil’

When the root it attaches to starts in ʁ-, the resulting consonant is not [ŋg] (after fortition) but [ŋk]:

(18)	[**ʁ**oleo]	‘to dry (in the sun)’	[ka-**ŋk**oleo-ha]	‘clothesline’
	[**ʁ**owea]	‘to carry’	[ka-**ŋk**owea-ha]	‘stretcher, litter’

Another context for nasal accretion is that of compounding, where the initial obstruent of the second word of a compound is fortified and prenasalized. Here we also see [ʁ] resulting in [ŋk].

(19)	[turu]	'drop'
	[**ʁ**ato]	'roof'
	[turu-**ŋk**ato]	'eaves'

As nasal accretion does not apply to sonorants (namely [l], [r] and [j]), one would not expect it to apply to [w]. Yet, [w] fortifies and prenasalizes to [mb] in compounds:

(20)	[aŋkaʔaŋka]	'follow'
	[**w**ono]	'smell'
	[no-aŋkaʔaŋka-**mb**ono-da]	'he followed them stealthily'

This, coupled with the aforementioned root-internal voicing constraint patterning, supports that [ʁ] and [w] operate underlyingly as obstruents, with the former being a voiceless dorsal and the latter being a voiced labial.

4.1.4 Banned word-finally

Kurmuk (Andersen, 2006, p. 60), Tigak (Beaumont, 1979, p. 15), Kensei Nsei (Akeriweh, 2000, p. 85) and Emmi (Ford, 1998, p. 43) all have a constraint banning final voiced obstruents. In line with this constraint, they all ban their unpaired voiced fricative(s) in syllable- or word-final positions.

4.2 Patterns only with sonorants

Table 3 lists all found sonorants and their representation in the database.

Table 3: distribution across languages of sonorant-patterning segments in dataset

Segment	n	Found in:
ʁ	4	Reunionnais, Portuguese, Cocos Malay, German
v	7	Mixtec, Korafe, Aghem, Skolt Saami, Kooma, Kawaiisu, Phlong
ɣ	6	Chukchi, Atayal, Aghem, Looma, Wambon, Kawaiisu
β	3	Atayal, Looma, Wambon
ʑ	1	Cofan
ʒ	3	Atayal, Kensei Nsei, Apinaye
ɮ	4	Tigak, Mongolian, Dongwang Tibetan, Ket
ʒ	1	Kensei Nsei
ʐ	4	Ronga, Tamil, Huichol, Armenian
ð^{j}	1	Scottish Gaelic
ʝ	1	Phlong

Below is an outline of the patterns observed within the set of consonants that show class behavior with sonorants.

4.2.1 Occupies sonorant positions in clusters

A very common way in which the sonorant status of a consonant becomes clear is through its distribution in consonant clusters. The most common of these positions is the second consonant of a bipartite complex onset. In both German (Orzechowska & Wiese, 2011) and Phlong (Cooke et al., 1976, p. 195), for instance, we see that C_1C_2 onsets are only allowed if C_2 is a sonorant.

(21)	German	[kʰlaɪn][1]	'small'
	Phlong	[blêʔ]	'pulverized'

In this position, German allows [ʁ] and both German and Phlong allow [v], revealing their sonorant status.

(22)	German	[kʰvaːl][1]	'agony'
		[kʰʁaŋk][1]	'ill'
	Phlong	[tve][2]	'to twist rope'

One also finds a similar patterning in final C_1C_2 clusters. Mongolian (Svantesson et al., 2005, p. 63), for instance, only allows such clusters if C_1 is a sonorant. Here we also see that [ɮ] occupies a sonorous position, despite its fricative-like noise and even tendency towards voicelessness.

(23)	[xɔŋx]	'bell'
	[art]	'people'
	[xɔɮp]	'to join'

The final pattern is represented by Mixtec (Hunter & Pike, 1969, p. 31), in which word-medial C_1C_2 clusters must have a sonorant occupying C_2 if C_1 is occupied by [ʔ]. Seeing as we observe [v] occupy C_2, in these clusters, we may conclude that [v] patterns with sonorants.

(24)	[kuʔni]	'to tie up'
	[koʔlo]	'turkey'
	[ʐaʔvi]	'expensive'

4.2.2 Occupies weak positions or alternates with consonant in a weakening environment

Closely related to the patterning in section 3.2.1 is one in which consonants join other sonorants in occupying weak positions, more specifically intervocalic and word-final positions. In Tigak (Beaumont, 1979, p. 15), word-final consonants may not be voiced obstruents. The fact that [ɮ] occurs in word-final position, then, shows that it is a sonorant.

[1] Normally, the aspiration of the k manifests as a devoicing of the following sonorant. For the sake of communication, this is not represented in these transcriptions.

[2] In fact, in this position, [v] alternates freely with [w], making it extra obvious as to its sonorance

The occurrence of these segments in weak positions is often the result of a lenition rule that operates on obstruents. In Wambon (Vries, 1992, p. 8), final [t] in intervocalic position lenites to [l ~ r], [p] to [β] and [k] to [ɣ].

(25)	[sit] + [-e]	>	[si**r**e ~ si**l**e]
	[ap] + [-e]	>	[a**β**e]
	[taɣimo] + [-ke]	>	[taɣimo**ɣ**e]

4.2.3 Forms pair with other sonorant

In Scottish Gaelic (Ladefoged et al., 1998, p. 3) and Dongwang (Bartee, 2007, p. 33), offending fricatives [ðʲ] and [ɮ] respectively form a systematic pair with a sonorants. In Dongwang, [ɮ] is the aspirated equivalent to [l], evidenced by the fact that [ɮ] causes breathy voice in the following vowel, as do all other aspirated sonorants (Dongwang forms aspiration pairs with glides and nasals). The authors of the source on Scottish Gaelic state that [ðʲ] is the palatalized counterpart to [r], but do not give language-internal synchronic evidence for that being the case. They only provide the information that some speakers of the Bernera dialect of Scottish Gaelic have [rʲ] in positions where the rest have [ðʲ].

4.2.4 Morphophonologically patterns or alternates with sonorant

Another common pattern within the dataset is one in which morphological context reveals an alternation with a sonorant. A known example is German (Antonsen, 2007, p. 31), whose rhotic [ʁ] alternates with the semivowel [ɐ] in pre-consonantal and word-final contexts.

(26)	[leːʁə]	‘learn’.1sg
	[leːɐt]	‘learn’.3sg

To give another example, Aghem (Hyman, 1979, p. 11) alternates vowels with their equivalent when two vowels come to border each other. In such a context, we can see that [ɣ] is treated as the glide counterpart of [a].

(27)	[a-] + [ɨn] > [ɣɨn]	cp.	[o-] + [ɨn] > [wɨn]

4.3 Straddlers

Tables 4 and 5 shows all found straddlers and their representation in the dataset with respect to their occurrence per language and their occurrence per language family and macroarea. The section hereafter lays out the straddling behavior in each language of the dataset.

Table 4: straddler distribution across languages in dataset

Segment	**n**	**Found in:**
ʁ	3	French, Northern Yukaghir, Southern Yukaghir
v	1	Apinaye
ɣ	5	Zapotec, Nchufie, Oykangand, Anguthimri, Ngan'gityemerri
ɬ	1	Chukchi
β	3	Zapotec, Anguthimri, Ngan'gityemerri
ʑ	2	Anguthimri, Ngan'gityemerri
ð	2	Zapotec, Anguthimri
z	2	Zapotec, Kurmuk (Burun)
g	1	Mongolian
ɢ	1	Mongolian
ʐ	3	Ngan'gityemerri, Zapotec, Mixtec

Table 5: straddler distribution across language families and macroareas

Macroarea	**Family**	**n**	**Languages**	**Segment(s)**
Eurasia	Indo-European	1	French	ʁ
	Yukaghir	2	Northern Yukaghir	ʁ
			Southern Yukaghir (Kolyma)	ʁ
	Mongolian	1	Mongolic	g, ɢ
North America	Otomanguean	1	Mixtec	ʐ
South America	Nuclear-Macro-Je	1	Apinaye	v
Africa	Atlantic-Congo	1	Nchufie (Bafanji)	ɣ
	Western Nilotic	1	Kurmuk (Burun)	z
North America	Otomanguean	1	Zapotec	ɣ, β, ð, z, ʐ
Australia	Pama-Nyungan	2	Oykangand	ɣ
			Anguthimri	ɣ, β, ð, ʑ
	Southern Daly	1	Ngan'gityemerri	z, ʐ, ɣ, β

4.3.1 French [ʁ]

The French rhotic [ʁ], orthographically represented as <r>, patterns with sonorants in at least two ways. Firstly, it occurs in the second position of an onset, a position only occupied by glides and liquids (Tranel, 1987, p. 137).

(28)	[pli]	'fold'
	[kʁi]	'shout'

Secondly, observe the two allomorphs of the morpheme in- 'not, un-' (Tranel, 1987, p. 77):

(29)	[ɛ̃-bʁyle]	'unburnt'
	[ɛ̃-paʁfɛ]	'imperfect'
	[ɛ̃-sasjabl]	'insatiable'
	[i-mɔʁtɛl]	'immortal'
	[i-lɔʒik]	'illogical'
	[i-nɔb̃ʁabl]	'innumerable'

Observing this, we see that the morpheme in- has the allophones [ɛ̃] and [i], the former being used before obstruents and the latter before sonorants. Given we see forms such as [i-ʁeɛl], we can conclude that [ʁ] in fact patterns with sonorants.

Despite this, we observe at least one situation in which [ʁ] patterns as an obstruent. Take the following vowel length distribution in word-final vowels (Tranel, 1987, p. 53).

(30a)	[bɛl]	'beautiful'
	[ʁɛn]	'reindeer'
	[gʁɛk]	'Greek'
	[gʁɛs]	'grease'
(30b)	[sɛ:z]	'sixteen'
	[bɛ:ʒ]	'beige'
	[ʁɛ:v]	'dream'
	[mɛ:ʁ]	'mother'

In (30b), we see that vowel lengthening is observed before the class [z, ʒ, v, ʁ], which consists of all voiced obstruents alongside [ʁ]. In this way, we see that [ʁ] patterns with voiced obstruents, and therefore obstruents in general.

In summary, French [ʁ] in fact does not only pattern with sonorants, as previously suggested in §2.2, but in fact patterns with both sonorants and obstruents, and is thus a straddler.

4.3.2 Yukaghir [ʁ]

In both Kolyma Yukagir (Grondona, 1998, p. 39), we observe a process of voice assimilation in obstruents, and not in sonorants.

(31a)	/leg-telle/	[lektelle]	‘eaten’
	/pad-telle/	[pattelle]	‘cooked’
(31b)	/qojl-pe/	[qojlpe]	‘gods’
	/irko:re-/	[irko:re]	‘frighten’

As expected of a voiced obstruent, [ʁ] also assimilates in voicing to the following consonant (additionally in matter, resulting here in a stop [q]), which places it among the obstruents.

(32)	/tʃoʁ-telle/	[tʃoqtelle]	‘cut’ (participle)

An additional process found in Yukaghir is word-final obstruent devoicing, as exemplified by the following words:

(33)	/pad/	[pat]	‘cook!’
	/leg/	[lek]	‘eat!’

Interestingly, however, we do not observe /ʁ/ devoicing word-finally. In fact, we observe it further leniting to [w], as the underlying form /tʃoʁ/ ‘cut!’ is realized as [tʃow]. This places it in the category of sonorants, which do not devoice word-finally.

4.3.3 Apinaye [v]

The phonotactics of Apinaye (Burgess & Ham, 1968, p. 7) show upon initial glance a fairly standard distribution of consonants in complex onsets. Complex onsets are allowed, as long as the first consonant is an obstruent, and the second is a sonorant. As such, as see words such as [ʔprõ] ‘his wife’ and [mro][1] ‘submerge’. Looking at the segment [v] in C_1C_2 onset clusters, we see that it occurs as both C_1 and C_2. See the following words:

(34a)	[ʔkvin]	‘broke it’
	[tvəm]	‘fat’
(34b)	[vrə]	‘descend’
	[vrəm]	‘empty’

Uniquely, however, [v] can occur in word-initial $C_1C_2C_3$ clusters as C_2, as exemplified by the following words:

(35)	[ŋvra]	‘buriti’
	[kvrəz]	‘parrot’

In this way, [v] represents a true phonotactic middle-point between obstruents and sonorants, as it is the only consonant allowed in this position. Phonetically, we see this segment allophonically alternate with [w] in coda position, making it a sonorant.

[1] Nasal stops pattern with oral stops in Apinaye, and in fact have prenasalized voiced plosive allophones in cluster-initial position.

4.3.4 Mongolian [g] and [ɢ]

The case of the Mongolian dorsal plosives [g] and [ɢ] is peculiar on multiple accounts.

On the phonetic level, they are the only voiced plosives in the language (Svantesson et al., 2005, p. 25), and in fact the only phonetically voiced obstruent in general, with the potential exception of the lateral fricative [ɮ], which patterns fully as a sonorant within the language (see §4.2.1). Additionally, they are the only plosives to not show a VOT contrast (Mongolian is an aspiration language, contrasting plain voiceless stops [p t] with aspirated equivalents [p^h t^h]). They additionally show no signs of devoicing and may occur in clusters among voiceless obstruents, implying that they are passively voiced. This is a property expected of sonorants and not plosives.

On the phonological level, [g] and [ɢ] show patterning of both sonorants and obstruents (Svantesson et al., 2005, p. 67). Firstly, they are allowed to participate in word-final C_1C_2 clusters as C_1, a position reserved for sonorants:

(36)	[ɢʊtə**m**tʃ]	'street'
	[sana**r**χ]	'to intend'
	[ta**w**s]	'salt'
	[ne:**g**t]	'to be opened'
	[ta**ɢ**t^h][1]	'balcony'

Inversely, they are allowed to be the second element of a C_1C_2 cluster if C_1 is a nasal, a position reserved for obstruents, as shown in (37).

(37)	[ɢʊrən**ts**]	'emery'
	[ʊn**t^h**]	'to sleep'
	[ʃitəm**s**]	'fillet'
	[ts^{h}ɔŋ**χ**]	'window'
	[m^{j}aŋ**ɢ**]	'thousand'
	[moŋ**g**]	'silver'

Thus, phonotactically, [g] and [ɢ] fill the positions of both sonorants and obstruents.

Finally, from a typological perspective, the presence of [g] and [ɢ] without the additional presence of [k] and [q] is highly unusual. Instances of a presence of a voiceless dorsal but the lack of a voiced one is commonly attested: Dutch has [k] but lacks [g], Bashkir has [q] but lacks [ɢ], and so on. Instances of voiced plosives lacking voiceless equivalents are typically reserved for labials, such as Arabic having [b] but lacking [p], and are virtually unattested for dorsals. Considering the possibility that the Mongolian dorsal stops actually function as sonorants within the language solves this problem, but leaves open the problem of the highly unusual mapping of dorsal plosives to underlying sonorants and the anomaly of allowing their participation in word-final CC clusters following a nasal.

4.3.5 Oykangand [ɣ]

Oykangand [ɣ] (Sommer, 1969, p. 43) patterns with sonorants in its alternation with [j] and [w] adjacent to high vowels. Take the suffix [-ɣ], expressing what in English is rendered using the modal auxiliary *should*. Before high vowels [i] and [u], it takes on the realisations [j] and [w]. This suggests a patterning (or at the very least demonstrates an alternation) with sonorants:

---

[1] Interestingly enough, [ɢ] can only occur together with [t^h] in such a cluster, rendering its distribution highly limited.

(38)	[elke**ɣ**] 'should return'	
	[ampi**j**] 'should taste'	
	[igu**w**] 'should go'	

On the other hand, we see an alternation with the obstruent [-g] in the dative suffix [-aɣ]. After a nasal, this suffix instead is realised as [-g], showing an alternation with obstruents.

(39)	[inaŋar**aɣ**] 'for (his) aunt'
	[olwon**g**] 'to the mountain'

Interestingly, in some words, the dorsal continuant is claimed to instead be voiceless, such as the word [ix] 'shell'. The source states this to still be a realisation of the same phoneme.

4.3.6 Nchufie [ɣ]

The dorsal continuant [ɣ] of Nchufie (Koopman & Kural, 1994, p. 3) shows an alternation with [k] (whereby intervocalic instances of /k/ are realised as [ɣ], neutralising the distinction between [k] and [ɣ]). Furthermore, [ɣ] may in fact sometimes surface as [g] instead. Additionally, we observe a patterning with [w], whereby, for instance, the singular agreement prefix /w-/ is realised as [ɣ-] in the presence of rounded back vowels (unrounded back vowels are present in Nchufie):

(40)	/w-o/	[ɣo]
	sgAgr-3PL	
	'their …'	

4.3.7 Anguthimri [ð], [ʑ], [β] and [ɣ]

The voiced 'fricative' series in Anguthimri (Crowley, 1981) has no voiceless equivalent. Given the Pama-Nyungan family is known for lacking fricatives altogether, there is cause to doubt that this voiced fricative series is a fricative series to begin with.

Firstly, we observe that these fricatives pattern with [r] in acquiring an epenthetic schwa in word-initial position, placing this series in a sonorant category (Crowley, 1981, p. 154):

(41)	/βaði/	'intestines'	[əβaði]
	/ðaj/	'mother'	[əðaj]
	/ʑoɣa/	'fly'	[əʑoɣa]
	/ɣama/	'child'	[əɣama]
	/ra/	'stomach'	[əra]

However, seeing as other sonorants such as [l], [ɹ], [j] and [w] do not acquire this epenthetic schwa, it is unclear what this rule is actually operating on. One could perhaps conclude here instead that [r] is a voiced obstruent, but that does not account for the onset cluster distribution discussed hereunder.

The phonotactics of Anguthimri, which allows complex onset clusters as long as they follow sonority sequencing, also allows [β] and [ɣ] alongside nasals to occupy initial positions of such clusters (Crowley, 1981, p. 155). The presence of nasals in such position is not too surprising, given we see frequently nasals pattern with other stops in such clusters (see the previously mentioned Apinaye but also Slavic clusters [ml-] and [nr-] such as Macedonian [mreʒa] 'net' and Russian [nravʲitsə] 'to like'). The presence of [β] and [ɣ] in those positions, implying they are stops, is however unexpected.

Typologically, Ngan'gityemerri is a language of the Pama-Nyungan family, a family known for its lack of fricatives. This makes the rich series of fricatives discussed in this section typologically unusual.

4.3.8 Kurmuk [z]

Kurmuk exhibits final obstruent devoicing. In Andersen (2007, p. 65), this is phrased in terms of a general ban of voiced obstruents in word-final position, with the addition of vowel-initial affixes “sometimes” causing intervocalic voicing (see [ɗɛ̀k] ‘tie’ vs. [ɗɛgu] ‘tied’). Given intervocalic voiceless obstruents are allowed in this language ([mɛtɪ] ‘to be beaten’), it is a more elegant analysis to treat these final segments as underlyingly voiced with a final obstruent devoicing rule rendering its realization as voiceless in word-final positions.

Kurmuk [z] is targeted by this devoicing rule ([ʔʊzʊ] ‘sucked’ vs [ʔʊʃ] ‘suck’), placing it in the category of obstruents. On the other hand, word-medial C_1C_2 clusters allow any sonorant to take the position of C_1 and any obstruent to take on the position of C_2. The segment [z] cannot take up either C_1 or C_2 (Andersen, 2007, p. 63). In this way, /z/ joins the sonorants in not being able to take up this position, but also joins the obstruents in not taking up C_1. This is, however, negative evidence and does not directly show alignment with either obstruent or sonorants.

4.3.9 Zapotec [ɣ], [β], [ð], [z] and [ʐ]

In Zapotec (Beam de Azcona, 2004, p. 46), word-final sonorants may receive glottalisation. The voiced fricative series also receives that glottalisation, placing them in the same category as sonorants (cp. examples 42a and 42b).

(42a)	[gæ̀jʔ]	‘five’
	[βrɛ̀lːʔ]	‘round’
	[sɣɑ̀mːʔ]	‘jicama’
(42b)	[mbæ̀zʔ]	‘fox’
	[ɟìʐʔ]	‘town’
	[stúβʔ]	‘another'
	[îðɟiðʔ]	‘sandal’
	[ɟìɣʔ]	‘frost’

In the specific case of [ɣ], we additionally observe an alternation with [j] and [w] when followed by a front and back vowel respectively.

Aside from the glottalisation applied to word-final sonorants, Zapotec has a separate so-called ‘glottal tone’ (Beam de Azcona, 2004, p. 61), which glottalises the vowel to which it is applied. After a glottal tone, a consonant of the voiced fricative series is devoiced alongside its glottalisation (see: /laʔz/ [laʔsʔ], /ɣuʔθ/ [ɣuʔθʔ], /ɟiʔβ/ [ɟiʔɸʔ], ...). As sonorants do not devoice under such circumstances, it may be argued that this is a rule that applies to the category of obstruents, placing the voiced fricatives in the obstruent category. Additional evidence for the obstruent analysis is provided by the fact that these segments are realized as stops (or in the case of [z] and [ʐ], affricates) when preceded by a homorganic nasal.

4.3.10 Ngan'gityemerri [ʐ], [ʑ], [ɣ] and [β]

Together with [b] and [d], the voiced fricatives [ʑ], [ɣ], [ʐ] and [β] of Ngan'gityemerri as described by Reid (1990, p. 152) are not allowed to occur word-finally, creating the class of voiced obstruents /b, d, ɣ, ʑ, ʐ, β/ . The individual sounds, however, all have their own peculiarities.

The sound labeled here as [ʐ] is described by Reid as having a 'rhotic release', transcribing it instead as [ʐɹ]. It additionally occurs in morphological formations when two instances of [r] come together, as in /ŋgir-ribem-tyalak/ 1pl-*stand-upright* [ŋgiʐɹibemtyalak] 'we stand up'. It also occasionally appears in free variation with [r], as in [mɑdɪwɪri] ~ [mɑdɪwɪʐɹi]. It patterns further with [r] and [ɹ] in being the only segments that do not occur word-initially. Finally, the sounds [ʐ ʑ r ɹ] all form a class in being the only sounds that cannot occur post-consonantally within a word-medial cluster. This supports that [ʐ] forms a class with [r] and [ɹ], presumably a class of rhotics. Whether [ʑ] belongs to this class is difficult to establish, as its patterning with the rest of the rhotics is limited and may be a statistical anomaly.

The consonant [ɣ] is found in word-medial rC clusters alongside [m, β and w]. This may suggest that [β] and [ɣ] belong together with [m] and [w] to form a class, or it could just be a statistical accident.

It is worth noting that [β], [ɣ] and [ʑ] have the tendency to be voiceless in word-initial positions, suggesting that they are obstruents. It is worth nothing, however, that this does not happen with the voiced plosives [b] and [d] (but does with [g]), despite the fact that they can phonetically afford to do so while maintaining their contrast with their voiceless counterparts, which are aspirated [pʰ] and [tʰ].

Typologically, Ngan'gityemerri is an Australian aboriginal language, an areal grouping known for its lack of fricatives. The presence of the fricatives discussed in this section is therefore typologically unusual.

4.3.11 Mixtec [ʐ]

As mentioned in §4.2.1, Mixtec word-medial C_1C_2 clusters must have a sonorant occupying C_2 if C_1 is occupied by [ʔ]. We used this to establish that [v] patterns with sonorants, as it can occupy this position C_2. Additional evidence for this comes from its realization [w] before [u] in this position. In line with this, we also find /ʐ/ in this position alongside [m n ɲ l v], placing it in a sonorant category.

(43)	[ku**ʔn**i]	'to tie up'
	[ɲa**ʔm**i]	'sweet potato'
	[ko**ʔl**o]	'turkey'
	[ʐa**ʔv**i]	'expensive'
	[ndiʃi**ʔʐ**u]	'goat'

Another kind of cluster allowed in Mixtec is NC where C is an obstruent and N is a homorganic nasal (Hunter & Pike, 1969, p. 31).[1] Here, we see [ʐ] occur alongside [d t s tʃ ʃ k], which places [ʐ] among the obstruents. It is worth noting that underlying /s/ and /ʐ/ surface as affricates post-nasally, which further strengthens the latter as a voiced obstruent.

(44)	[**nd**aʔa] 'hand'
	[**nts**oo] 'to be carrying'
	[**ŋk**eesã] 'I sank'
	[**ndʐ**oso] 'was mounted'

[1] Whether these are clusters or represent a prenasalized obstruent series is unclear. The author refers to them as clusters.

4.4 Elusive cases

Below are cases found that are difficult to categorize under any of the categorizations presented in the previous sections.

4.4.1 Kadiweu [ʁ]

In Kadiweu, [ʁ] varies freely with [ɢ] in word-initial position (Sandalo, 1995, p. 18). Given this, and that the language has [q] in addition, one may conclude that [ʁ] is the voiced equivalent to [q]. This would put it in the same category of the consonants of §4.1.1.

We also observe that [ʁ] is allowed in C_1C_2 onset clusters as C_2 (Sandalo, 1995, p. 20). In fact, [ʁ] is the only consonant allowed in this position. This may imply that [ʁ] is some sort of rhotic, given rhotics are typical occupants of the C_2 position within onset clusters. Ultimately, however, this is too little information to make a conclusion on its classification.

4.4.2 Chukchi [ɬ]

The lateral fricative of Chukchi forms a natural class with the sound /s/ in forming an environment for the denasalization of /ŋ/ right after it (Dunn & others, 1999, p. 46). Thus, the sequence /ɬŋ/ is rendered [ɬɣ]. See the example /**ɬəŋ**ək/ [**ɬəŋ**ək] > /nə**ɬŋ**əqin/ [nə**ɬɣ**əqin]. As such, we see /ɬ/ belonging together with /s/ to form a class of voiceless fricatives.

On the flipside, observing the phonotactics of the language shows us that /ɬ/ forms a natural class additionally with /r/ alongside /s/ in that normally obligatory word-initial inter-cluster schwa-insertion is optional in an underlying occurrence of /Cɬ/ (Dunn & others, 1999,). Thus, the sequences /#pɬ/, /#pr/ and /#ps/ all may surface as [pɬ], [pr] and [ps] respectively, against the usual rule that resolves word-initial /CC/ clusters through schwa insertion [CəC].

In the Chukchi case, thus, we see a patterning of [ɬ] with both [s] and [r], rendering the underlying sonorance of [ɬ] ambiguous.

4.5 Summary of patterns

Below is a summary of all offending segments found and the patterns observed in their class behavior.

4.5.1 Summary of obstruent patterns

Within consonants found to pattern with obstruents only, the following patterns were observed:

1. The consonant in question alternated with and/or formed a distributional class with stops. Ex. Mocovi /ʁ/ which is realized [q] (merging with /q/) word-finally and [ʁ] elsewhere.
2. The consonant in question shows free variation in voicing, or alternates as if it is voiceless. Ex. Mavea [β], which alternates freely with [ɸ].
3. The consonant is not allowed to occur word-finally, just like other voiced obstruents.

4.5.2 Summary of sonorant patterns

Within consonants found to pattern with sonorants only, the following patterns were observed:

1. The consonant occupies a position in clusters only reserved for sonorants. Ex. Mongolian [ɮ], which is allowed in C_1 position in C_1C_2 clusters, a position researved for only.
2. The consonant occurs in a weak position alongside other sonorants, or is the result of a weakening process. Ex. In Wambon, [ɣ] is the result of an intervocalic lenition process that otherwise results in sonorants. Thus, we see [t] > [r ~ l], and similarly [k] > [ɣ].

3. The consonant forms a symmetric pair with a sonorant within the system of the language. Ex. In Scottish Gaelic, [ðʲ] is the palatal counterpart of [r].
4. Morphophonologically alternates with a sonorant. Ex. in Aghem, [ɣ] is the result of a gliding process targetting [a] before another vowel, the some process that targets [o] to result in [w].

4.5.3 Summary of straddler patterns

The straddlers found in this paper have been presented in the form of individual case studies, in order to ensure a descriptive depth worthy of its focus within the research. Here, I present a summary of patterns spotted within those case studies:

1. Passive role: a common pattern observed with straddlers is that while they show class behavior with obstruents as undergoers of processes rather than as causes for change, in which their role would be more active. For instance, they may undergo assimilation, but never be assimilators themselves. Kurmuk [z] devoices word-finally; Yukaghir [ʁ] undergoes voice assimilation, but does not cause it itself.
2. Alternation with both a sonorant and an obstruent: multiple straddlers within this study show alternations both with a sonorant and with an obstruent. Oykangand [ɣ] alternates with [g] after nasals, but [j] word-finally before a front vowel. Nchufie [ɣ] varies freely [g], but shows morphological alternations with [w] in word-initial positions. Yukaghir [ʁ] alternates with [q] before voiceless plosives, but with [w] word-finally.
3. Shows the prosodic distribution of both a sonorant and an obstruent: straddlers can also be seen to occupy prosodic positions reserved for obstruents and sonorants independently. Mixtec [ʒ] occupies sonorant positions in word-medial clusters, but occupies obstruent positions in word-initial clusters. Mongolian [g] and [ɢ] occur in final C_1C_2 clusters as either C_1 or C_2. The former is otherwise reserved for sonorants, while the latter is reserved for obstruents.

5. Discussion

The hypothesis of the present research is that many voiced obstruents with no voiceless counterpart will display phonological class behavior with sonorants and obstruents separately. In the previous section, I have described the phonological behavior patterns of 102 unpaired voiced obstruents (primarily fricatives) across 58 languages. The phonological class behavior of 25 of them could not be established. Out of the remaining 77, almost a third (24 spanning 12 languages) showed separate class behavior with both sonorants and obstruents, in line with this hypothesis.

The data presented in this paper challenges sonorance from multiple angles. Firstly, it shows that phonetic cues – at least presuming that they are accurately described by the sourced authors – are unreliable in establishing the sonorant status of a given sound. Secondly, it shows that a strict binary view of sonorance as a phonological property, which is the view adopted by most feature theories, is not borne out in cross-linguistically observed phonological patterning. While this study has shown apparent cases of segments that go against these notions, it excludes theoretical accounts from its scope. This will be tackled separately in following research. The solution to these problems may lie within theories in which sonorance is not an independent category but is instead an informal grouping of sounds that happen to be of a certain form. For instance, element[1]-based approaches provide components that represent properties typically associated with obstruents (such as |H|, representing representing high frequency energy typically associated with that of fricatives) but do not formally discriminate between sonorants and obstruents (Backley, 2011, p. 130).

[1] An 'element' being a privative component corresponding to resonance properties in the speech signal.

The scope of this research, as affected by temporal constraints, has been constrained to voiced fricatives selected on the basis of their lack of pairing with a voiceless counterpart. This significantly reduces the final pool of potential obstruent-sonorant straddlers, especially given that they do not necessarily have to meet the criteria outlined in this paper. Firstly, straddlers do not have to be fricatives, as exemplified by Mongolian, in which they take the form of dorsal plosives. Secondly, they do not have to be unpaired, as seen in the two examples given in §2.3: Russian /v/ does on the surface seem have a voiceless counterpart /f/, and Flemish Dutch /ɣ/ has a voiceless counterpart /x/. Future research may expand the pool to include voiced plosives and/or voiced fricatives regardless of their pairing, although this may require better selection criteria, more accurate methods of filtering out information from the current publicly-available phonological databases and, most important of all, fieldwork that places a greater focus on phonology.

It is further worth noting that straddling behavior may occur across domains other than the sonorance. For instance, §4.1.1 shows examples of phonemes that instead straddle the line between plosive and fricative. One may also observe potential straddling behavior in nasals, who are traditionally viewed as sonorants but may also appear in skeletal positions typical of obstruents (see discussion in §4.3.7).

6. Conclusion

In conclusion, the concept of sonorance, especially presented as a binary distinction, does not yield reliable phonetic or phonological predictions in its current mainstream conceptualization. On the aspect of the phonetics-phonology interface, it does not reliably predict the phonological behavior of a sound based on its phonetic manifestation. From a phonological perspective, a segmental unit may share behavioral patterns of both sonorous and obstruent consonants independently, a situation directly excluded by the conceptualization of sonorants and obstruents as two opposites of the same coin. A reconceptualization of sonorance is necessary to account for this, either by adopting a less rigid notion of sonorance, or by dispensing of it altogether in formal representation.

Bibliography

Akeriweh. (2000). *A Step Towards the Standardisation of K`answéynséy (A Grassfield Bantu Language)* [Master's Thesis]. Universite de Yaounde I.

Andersen, H. (1969). The phonological status of the Russian 'labial fricatives.' *Journal of Linguistics*, *5*(1), 121–127.

Andersen, T. (2006). Kurmuk phonology. *Studies in African Linguistics*, *35*(2), 29–90.

Antonsen, E. H. (2007). Elements of German: Phonology and morphology. In *Elements of German: Phonology and morphology*. University of Alabama Press.

Backley, P. (2011). *Introduction to element theory*. Edinburgh University Press.

Bartee, E. L. (2007). *A grammar of Dongwang Tibetan* [PhD Thesis]. University of California, Santa Barbara.

Beam de Azcona, R. G. (2004). *A Coatlan-Loxicha Zapotec Grammar* [PhD Thesis]. University of California, Berkeley.

Beaumont, C. H. (1979). *The Tigak Language of New Ireland* (Vol. 58). Australian National University.

Blaho, S. (2008). *The syntax of phonology: A radically substance-free approach*.

Booij, G. (1999). *The phonology of Dutch*. Oxford University Press.

Botma, B., & van't Veer, M. (2013). A fraction too much friction: The phonological status of voiced fricatives. *Linguistics in the Netherlands*, *30*(1), 46–60.

Burgess, E., & Ham, P. (1968). Multilevel Conditioning of Phoneme Variants in Apinaye. In *Linguistics* (Vol. 41, pp. 5–18).

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology*, 283–333.

Coats, H. S., & Harshenin, A. P. (1971). On the phonological properties of Russian v. *The Slavic and East European Journal*, *15*(4), 466–478.

Cooke, J. R., Hudspith, J. E., & Morris, J. A. (1976). Phlong (Pwo Karen of Hot District, Chiang Mai). In W. A. Smalley (Ed.), *Phonemes and Orthography: Language Planning in Ten Minority Languages of Thailand* (Vol. 43, pp. 187–220). Australian National University.

Crowley, T. (1981). The Mpakwithi Dialect of Anguthimri. In R. M. W. Dixon & B. Blake (Eds.), *Handbook of Australian languages* (Vol. 2, pp. 147–194). Australian National University Press.

Dunn, M. J. & others. (1999). *A grammar of Chukchi*.

Ford, L. J. (1998). *A Description of the Emmi language of the Northern Territory of Australia* [Phdthesis, The Australian National University]. http://hdl.handle.net/1885/10796

Friedman, V. A. (2002). *Macedonian*. Lincom Europa.

Grondona, V. M. (1998). *A Grammar of Mocovi* [PhD Thesis]. University of Pittsburgh.

Guerin, V., & Aoyama, K. (2009). Illustrations of the IPA: Mavea. *Journal of the International Phonetic Association*, *39*(2), 249–262.

Hunter, G. G., & Pike, E. V. (1969). The Phonology and Tone Sandhi of Molinos Mixtec. *Linguistics*, *47*, 24–40.

Hyman, L. (1979). *Aghem Grammatical Structure*. University of Southern California.

Iosad, P. (2018). *There is no problem of/v/in Russian phonology*.

Koopman, H., & Kural, M. (1994). *Aspects of Nchufie grammar*. UCLA.

Ladefoged, P., Ladefoged, J., Turk, A., Hind, K., & others. (1998). Phonetic structures of Scottish Gaelic. *Journal of the International Phonetic Association*, *28*(1–2), 1–41.

Lulich, S. (2004). *Russian [v]: An acoustic study*.

Maslova, E. (2003). *A Grammar of Kolyma Yukaghir*. Mouton de Gruyter.

Mielke, J. (2005). Emergent feature theory. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *41*(2), 259–274.

Mielke, J. (2008). *The emergence of distinctive features*. Oxford University Press.

Moran, S., & McCloy, D. (Eds.). (2019). *PHOIBLE 2.0*. Max Planck Institute for the Science of Human History. https://phoible.org/

Morgenstierne, G. (1945). Notes on Burushaski Phonology. *Norsk Tidsskrift for Sprogvidenskap*, *13*, 61–95.

Orzechowska, P., & Wiese, R. (2011). Reconstructing the Sonority Hierarchy. *ICPhS*, 1542–1545.

Padgett, J. (2002). Russian voicing assimilation, final devoicing, and the problem of [v]. *Natural Language and Linguistic Theory*.

Petrova, O., & Szentgyörgyi, S. (2004). */V/and voice assimilation in Hungarian and Russian*.

Reid, N. J. (1990). *Ngan'gityemerri: A Language of the Daly River Region, Northern territory of Australia* [Phdthesis]. The Australian National University.

Reiss, C. (2017). Substance free phonology. In *The Routledge handbook of phonological theory* (pp. 425–452). Routledge.

Sandalo, M. F. (1995). *A Grammar of Kadiweu* [PhD Thesis]. University of Pittsburgh.

Sibanda, G. (2004). *Verbal phonology and morphology of Ndebele* [PhD Thesis]. UC-Berkeley.

Sommer, B. A. (1969). *Kunjen phonology: Synchronic and diachronic*. Pacific Linguistics.

Svantesson, J.-O., Tsendina, A., Karlsson, A., & Franzén, V. (2005). *The phonology of Mongolian*. OUP Oxford.

Tolkov, I. (2011). *Phonetics of Sakha*.

Tranel, B. (1987). *The sounds of French: An introduction*. Cambridge university press.

van den Berg, R. (1989). *A Grammar of the Muna Language*. Foris Publications.

Van Oostendorp, M. (2003). Ambisyllabicity and fricative voicing in West Germanic dialects. *The Syllable in Optimality Theory*, 304–337.

Vries, L. de. (1992). *The Morphology of Wambon of the Irian Jaya Upper-Digul Area*. KITLV Press.

Wilbur, J. (2014). *A Grammar of Pite Saami*. Language Science Press.

Appendix – Excluded languages

Excluded language	**Reason**
Hill Mari	Source could not be retreived in language understood by the author
Inuit	Source could not be retreived in language understood by the author
Mazatec	Source could not be retreived in language understood by the author
Nganasan	Source could not be retreived in language understood by the author
San Bartolome Zapotec	Source could not be retreived in language understood by the author
Lebanese Arabic	Source could not be retreived in language understood by the author
Northern Altai	Source could not be retreived in language understood by the author
Eastern Khanty	Source could not be retreived in language understood by the author
Campidanese Sardinian	Source could not be retreived in language understood by the author
Amele	Transcription incorrectly listed on Phoible
Aleut	Source does not incluide relevant information
Nyinpa Cone	False positive
Dongxiang	Source does not incluide relevant information
Nuosu	False positive
Leggbo	Source does not incluide relevant information
Awngi	Transcription incorrectly listed on Phoible
Saotomense	False positive
Vame	Source does not incluide relevant information
Fang	Source does not incluide relevant information
Sui	Source does not incluide relevant information
Linggithigh	Source does not incluide relevant information
Khalkhal	Source does not incluide relevant information
Lagwan	False positive
Buharan Tajik	Source does not incluide relevant information
Italian	Source positive
Guajarara	Transcription incorrectly listed on Phoible