Phonotactic Evidence from Typology and Acquisition for a Coda+Onset Analysis of Initial sC Clusters

Heather Goad

1. Introduction

Recent research in phonology has seen structural and perceptual approaches to syllable well-formedness pitted against each other. Concerning left-edge clusters, for example, different explanations are provided for the observation that obstruent-initial (*play*) and *s*-initial clusters (*stay, slay*) diverge in phonological behaviour.[1] Researchers who advocate a structural approach propose that this motivates an articulated view of the syllable: obstruent-initial and *s*-initial clusters have different representations (see Goad 2011 for a recent overview). Researchers who advocate a perceptual approach, by contrast, propose that these differences can be explained solely by perceptual considerations: segments are ordered to maximize their perceptibility and the acoustic properties of *s* versus other obstruents account for their different behaviour (e.g. Wright 2004). In this paper, I return to a structural approach to syllable well-formedness, but one which is informed by perceptual considerations.

In view of the perspective I take, one question that arises is what type of structure holds for word-initial sC clusters, given the multitude of options that have been proposed. As will be shown, initial sC clusters mirror the phonotactic profile of ordinary coda+onset clusters; for instance, *s*+stop is always well-formed, while *s*+sonorant is not (Goad 2011, 2012). Because of this, I adopt the position of Government Phonology that sC clusters are analysed as coda+onset clusters (following Kaye 1992) and, critically, that this position holds for the left edge of words, as well as medially.

In the latter half of the paper, I turn to explore the predictions that stem from a coda+onset analysis of sC clusters for second language (L2) acquisition. Specifically, it is predicted that in the acquisition of a subset grammar (i.e., where the L2 is contained within the L1 for the construction under focus), learners exposed only to the ill-formedness of *s*+stop should infer the ill-formedness of *s*+sonorant. A related question that emerges in this context concerns the kinds of evidence that learners can use to acquire subset grammars. I present some results from Schwartz & Goad (2015) who experimentally show that the acquisition of a subset grammar can be facilitated by exposure to a new type of evidence that has not, to our knowledge, been explored in L2 acquisition: Indirect Positive Evidence (IPE). IPE is evidence from errors in the learner's L1 made by native speakers of the learner's L2.

2. Phonotactics

I begin by examining the role that sonority plays in syllable well-formedness. It has long been known that segments are optimally ordered to ensure that there is a (steep) rise in sonority toward the nucleus, a (shallow) fall in sonority away from the nucleus, and a fall in sonority across a syllable boundary. More concretely, the consequences for word-initial clusters are that: (i) they strive to be

[1] Throughout the paper, 'obstruent' refers to stops and fricatives other than *s*, except where noted.

obstruent-initial; and (ii) steep sonority rises are preferred over shallow rises, which are in turn preferred over sonority plateaus. These observations, which can be formally expressed as a set of constraints on cluster well-formedness, are reflected in the typology of obstruent-initial clusters shown in (1).

(1) Typology of obstruent-initial clusters (word-initial position):

	English	Dutch	Greek
obstruent+liquid	✓	✓	✓
obstruent+nasal	*	✓	✓
obstruent+obstruent	*	*	✓

sC clusters, by contrast, optimally show a sonority plateau. As the sonority of C in sC rises, the cluster worsens (Goad 2011, 2012). This is shown in (2).[2]

(2) Typology of *s*-initial clusters (word-initial position):

	Brazilian Portuguese	French	Greek	Dutch	English
s+stop	*	✓	✓	✓	✓
s+nasal	*	*	(✓)	✓	✓
s+lateral	*	*	*	✓	✓
s+rhotic	*	*	*	*	✓

In short, sC clusters have the inverse profile of obstruent-initial clusters on the sonority dimension.

3. Structural approaches to the syllable

In structural approaches to the syllable, the observation that *s*+stop clusters violate the constraint that obstruent-initial clusters rise in sonority has motivated different representations for both cluster types. Obstruent-initial clusters form left-headed branching onsets (e.g. Kaye, Lowenstamm & Vergnaud 1990) (3a), while *s* is located outside the onset constituent in sC clusters, regardless of (at least for most researchers) the sonority profile of the sC cluster. Exactly how *s* is organized in the latter has been the subject of much debate (for alternatives, see e.g. Steriade 1982, van der Hulst 1984, Goldsmith 1990, Kaye 1992). (3b) suffices for the moment but we will return to address this question more concretely in §6.

(3) a. Obstruent-initial cluster: b.   sC cluster (minimal representation):

```
        O                                         O
       /|\                                         |
      / | \                                        |
     p    l                                   s    p
                                              s    l
```

In approaches along the lines of (3), phonological units are highly articulated and, thus, the burden of explanation is placed on the structural relationships that adjacent segments enter into. Although this type of approach can thereby capture many of the peculiarities that sC clusters display vis-à-vis obstruent-initial clusters, an important question that remains unanswered in this literature is: why is the consonant that displays unorthodox behaviour always *s* (and/or other sibilants)? This question is answered in perceptual approaches to the syllable, which we turn to next.

---

[2] The parentheses indicate that *s*+nasal is marginally productive in Greek. *s*+rhotic in English is realized as [ʃr].

4. Perceptual approaches to the syllable

A growing literature has moved away from the position that syllables are highly articulated units along the lines of (3) (e.g. Steriade 1999, Côté 2000, Fleischhacker 2001). Segments are instead proposed to be ordered to maximize their perceptibility (e.g. Wright 1996, 2004). Word-initial cluster well-formedness reflects this: obstruents (especially stops) have weak internal cues; they are optimally identified in contexts where they can 'lean on' a following sonorant, thereby explaining the preference for a rising sonority profile in left-edge clusters (✓tr, ✓kl). Unlike other obstruents, strident fricatives have robust internal cues, ensuring that their place and manner can be appropriately identified even in non-optimal contexts, which accounts for the well-formedness of s+stop (✓sp, ✓st).

If the acoustic properties of *s* vs. other obstruents can explain cluster well-formedness, we must question whether the structural approach to syllable phonotactics can be dispensed with altogether in favour of a purely perceptual approach. Indeed, there are patterns of behaviour displayed by left-edge clusters that appear not to be well handled by the structural approach outlined in (3). We consider in this respect preferred epenthesis site in languages that repair both sC and obstruent-initial clusters. Some questions that arise that seem to challenge the structural approach are as follows (see further §5). (i) Why is prothesis preferred for some types of clusters (e.g. *sta → əsta*), while anaptyxis is preferred for others (e.g. *tra → təra*)? (ii) Why do sC clusters not necessarily all pattern together in a language (e.g. *sna → əsna* but *sla → səla*)?
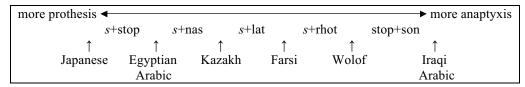
In Fleischhacker's (2001) examination of epenthesis site in L2 acquisition and loanword adaptation, she aims to answer questions such as these. She proposes that the epenthesis site strives to optimize perceptual similarity between input and output.[3] That is, it occurs in a location where it is least conspicuous, which depends on the particulars of the segments contained in the cluster. The specific perceptually-based predictions that Fleischhacker makes are in (4) (where *ə* represents the epenthetic vowel):

(4) Fleischhacker's predictions:
1. Anaptyxis over prothesis in stop+sonorant (*təra > ətra*)
2. Prothesis over anaptyxis in *s*+stop (*əsta > səta*)
3. Among *s*+sonorant sequences, more anaptyxis as the sonorant increases in sonority (*əsna* but *səla*)
4. More anaptyxis in stop+sonorant than in fricative+sonorant (*təra > fəra/səra*)

The predictions in (4) translate into a set of constraints, which interact with other constraints in a language to yield the typology of epenthesis sites in Figure 1.

Figure 1. Typology of epenthesis sites (adapted from Fleischhacker 2001).

more prothesis ◄────────────────────────────► more anaptyxis					
s+stop	s+nas	s+lat	s+rhot	stop+son	
↑	↑	↑	↑	↑	↑
Japanese	Egyptian Arabic	Kazakh	Farsi	Wolof	Iraqi Arabic

Languages like Japanese, at one end, exclusively employ anaptyxis to repair left-edge clusters (due to overriding coda constraints), while languages like Iraqi Arabic, at the other end, exclusively employ prothesis (due to the requirement that input and output segments be contiguous). Other languages fall between these two extremes, depending on the sonority value of the consonant following *s* or whether the cluster begins with *s* or a stop. Figure 1 visually illustrates that Predictions 1-3 are supported (once

---

[3] The input is the surface form in the target language in L2 acquisition and the surface form in the source language in loanword adaptation. The output is the surface form arising from the L2er's grammar and the surface form in the borrowing language in loanword adaptation.

other constraints are factored in). Prediction 4, which is not evident from Figure 1, poses more of a challenge.

To further explore Prediction 4, let us consider the predictions of a purely perceptual approach in contrast to a purely structural approach. Fleischhacker's approach does not predict that strident and non-strident fricatives should pattern differently when they are followed by liquids (or other consonants that are high in sonority). Thus, strident fricative + liquid should pattern identically to non-strident-fricative + liquid in how the cluster is repaired (*sl = fl*). The structural approach, by contrast, predicts that these two cluster types should pattern differently, as all sC clusters, even those that rise in sonority, are standardly assumed to have a different representation from rising sonority clusters that begin with other fricatives (*sl ≠ fl*). Non-strident fricative-initial clusters should instead pattern with stop-initial clusters (*fl = pl*). The data below on epenthesis site from Farsi-speaking learners of English support the structural approach: non-strident fricative + sonorant (in (5b)) patterns with stop + sonorant (5a), and not with *s* + sonorant (5c), which patterns with *s* + stop (5d).

(5) L2 Farsi English (Karimi 1987):
   a. p[e]lastic   b. f[i]loor   c. [e]smoke   d. [e]sp[i]ring
       p[e]roud   F[e]red   [e]snow   [e]statistic
       d[i]rink   th[i]ree   [e]slide   [e]ski (loan)

5. Challenges

To summarize thus far, we have arrived at one set of challenges for a purely perceptual approach to syllable well-formedness and another set of challenges for a purely structural approach to syllable well-formedness. These are provided in (6) and (7) below:

(6) Challenges for a perceptual approach:
- Why do *s*+sonorant and fricative+sonorant clusters pattern differently, when both display the same rising profile?
- Why does no language that permits sC clusters forbid *s*+stop? Although a perceptual approach can account for the well-formedness of *s*+stop, it is not expected to be preferred among sC clusters.
- Why does the well-formedness of sC deteriorate as the sonority of C increases, the inverse profile of obstruent-initial clusters?

(7) Challenges for a structural approach:
- Why is it *s* that behaves as special in languages?
- Why is prothesis preferred as a repair for sC clusters and why is anaptyxis preferred as a repair for obstruent-initial clusters? Although these two types of clusters are not expected to behave identically (compare their different representations in (3)), it is not evident why it is sC that prefers prothesis and obstruent-initial that prefers anaptyxis.
- Why do sC clusters with higher sonority C prefer anaptyxis while sC clusters with lower sonority C prefer prothesis? Given the representation in (3b), all sC clusters should behave in the same manner.

As is evident, the challenges for each of these approaches are complementary and, thus, we are confronted with the task of finding a solution where both sets of challenges can be resolved. With this goal in mind, I suggest that perceptual factors have the following responsibilities for left-edge cluster well-formedness:

(8) Responsibilities of perceptual factors:
- Segments are ordered to maximize their perceptibility, modulo overriding structural constraints.
- As sonority is acoustically-defined (minimally, as relative intensity), segments that are more intense are more perceptible. Left-edge clusters optimally rise in sonority, modulo overriding

structural constraints, because obstruents rely on following sonorants to ensure that they are appropriately perceived (i.e., perceived as distinct from other segments with which they may be confused).

- Because of its acoustic properties, *s* (and sibilants more generally) need not be followed by consonants of higher sonority; *s* has strong internal cues to place and manner, which positively impacts its perceptibility (i.e., ensures that it is appropriately identified regardless of what it is adjacent to).

The rest of left-edge cluster well-formedness lies to structure in a way that will be articulated in the following section. We begin §6 by probing the representation of sC clusters, to come to some understanding of the overriding structural constraints referred to in (8).

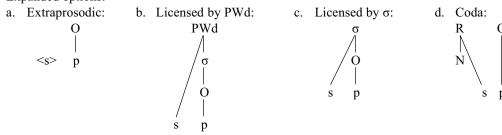## 6. What is the right analysis for sC?

In §3, a minimal representation was provided for left-edge sC clusters, repeated here as (9). At that point, all that was necessary was to ensure that *s* be located outside the constituent that organizes the following onset. In reality, several options have been proposed for *s*, all of which expand on (9).

(9) Minimal representation (from (3b)):

O
|
s   p

The most commonly motivated options are provided in (10) (see Goad 2011 for further discussion of alternatives). In (10a), *s* is extraprosodic (e.g. Steriade 1982): it is not organized into higher structure but is protected from deletion. In (10b) and (10c), *s* is organized into higher structure, the prosodic word (e.g. Goldsmith 1990) and syllable (van der Hulst 1984) respectively.[4] In (10d), *s* is the coda (technically, rhymal dependent) of an empty-headed syllable (following Kaye 1992). (The representation in (10d) is a simplified version of that adopted by Kaye and government phonologists more generally.)

(10) Expanded options:
  a. Extraprosodic: b. Licensed by PWd: c. Licensed by σ: d. Coda:

To tease apart these four options, we consider the predictions that each makes concerning the phonotactic constraints that *s* enters into with the following consonant. Under options (10a-c), the constraints that hold between *s* and C should not mirror those that hold of segments in any other type of cluster (branching onset, coda+onset, etc.). This is unsupported. Under option (10d), the phonotactic constraints that hold between *s* and C should mirror those that hold of segments in coda+onset clusters. This is supported.

To motivate the claim that initial *s*C clusters mirror the phonotactic constraints that hold of coda+onset clusters, we return to the typology of sC cluster well-formedness introduced in (2):

---

[4] Exactly which constituent licenses *s*, PWd or σ, is likely parametrically determined, depending on the type of domain at which left-edge sC is permitted in a given language (Goad & Rose 2004, Goad 2012). See also Vaux & Wolfe (2009).

(11) Typology of *s*-initial clusters (word-initial position) (repeated from (2)):

	Brazilian Portuguese	French	Greek	Dutch	English
s+stop	*	✓	✓	✓	✓
s+nasal	*	*	(✓)	✓	✓
s+lateral	*	*	*	✓	✓
s+rhotic	*	*	*	*	✓

As was mentioned earlier, (2)/(11) shows that sC clusters have the inverse profile of branching onsets on the sonority dimension: as the sonority of C increases, the cluster worsens. What, though, is the evidence that these clusters mirror the phonotactic constraints that hold of coda+onset clusters? For this, we turn to a discussion of syllable contact.

The Syllable Contact Law, from the seminal work of Murray & Vennemann (1983), is provided in (12):

(12) Syllable Contact Law (Murray & Vennemann 1983: 520):
The preference for a syllabic structure *A$B*, where *A* and *B* are marginal segments and *a* and *b* are the Consonantal Strength values of *A* and *B* respectively, increases with the value of *b* minus *a*.

Consonantal strength is the inverse of sonority. The sonority scale I adopt is: obstruent (including *s*) < nasal < lateral < rhotic. This sonority scale is standard (e.g. Clements 1990), except for the finer division drawn between types of liquids. The higher sonority of rhotics over laterals is motivated by the difference between *s*+lateral and *s*+rhotic in (11) and appears quite extensively in the literature as well (e.g. Selkirk 1984, Hall 1992, Smith 2005).[5] Using this sonority scale, we arrive at the consonantal strength values in (13):

(13)

Sonority scale:	obstruent < nasal < lateral < rhotic
Consonantal strength:	4 3 2 1

Let us consider word-internal coda+onset well-formedness, when the coda is an obstruent (including *s*), from the perspective of (12) and (13). Consider (14).

(14) Word-internal position:
Segmental profile:	Vp.tV	>	Vp.nV	>	Vp.lV	>	Vp.rV
	Vs.tV	>	Vs.nV	>	Vs.lV	>	Vs.rV
Syllable contact (*b-a*):	0		-1		-2		-3

As can be seen, as sonority increases throughout the cluster, the coda+onset parse becomes less optimal due to poor syllable contact. Indeed, once a rising sonority profile reaches some threshold, word-internal clusters prefer to be syllabified as branching onsets. This yields, for instance, *V.plV/V.prV* over **Vp.lV/Vp.rV*. However, if the branching onset option is never available for sC clusters (Goad 2012),[6] only *Vs.lV/Vs.rV* is possible. In many languages, the poor sonority profile of the

---

[5] See Parker (2002) for an extensive review of sonority, including sonority scales.

[6] Ashley Farris-Trimble has brought to my attention a handful of understudied languages in which the only sC cluster is *s*+liquid in shape. As obstruent+liquid is the optimal branching onset, it behooves us to ask whether in these languages *s*+liquid forms a branching onset. I hypothesize that it does, but that in these languages, *s* may be low in stridency; this would make it ineligible to participate in true sC (coda+onset) clusters, given the need for *s*+stop to be well-formed in all languages with sC, precisely the context where the perceptibility of *s* would be most compromised. It stems from this hypothesis that *sl/sr* clusters in these languages would be more akin to *f*- or *θ*-initial clusters. For further discussion of the role of stridency in the unusual behaviour of *s*, see Goad (2012, to appear).

coda+onset parse rules these clusters out altogether (as in, for example, French: *[asʁɛ], cf. [atʁɛ] 'attraction'; *[tasle], cf. [takle] 'to tackle').

Turning to word-initial position, if sC clusters in this position also form coda+onset clusters, the same effect will be observed: *.plV/.prV* over *Øp.lV/Øp.rV* (where Ø represents an empty nucleus) but *Øs.lV/Øs.rV* over **.slV/.srV*). Although the poor syllable contact in *Øs.lV/Øs.rV* is tolerated in some languages (e.g. English: *sled/shred*; see (11)), as the sonority profile of sC increases, the well-formedness of the cluster gets worse and is ruled out in other languages (e.g. French: *[srɛ], cf. [tʁɛ] 'very'; *[sla], cf. [pla] 'flat'; see (11)).[7]

Although we have seen that a coda+onset analysis of sC clusters, coupled with syllable contact, yields the attested typology in (11), we must return to Fleischhacker's (2001) results on preferred epenthesis site in cluster repair to see if the structural approach taken here can account for the patterns observed. In the following lines, we provide structural explanations for Fleischhacker's Predictions 1 to 3 (from (4)), based on syllable contact (see earlier Gouskova 2001). (Prediction 4 was previously addressed and supported from a structural perspective, in §4 above.) Prediction 1 stated that anaptyxis is preferred over prothesis in stop+sonorant. From a structural perspective, anaptyxis is favoured for stop+sonorant because heterosyllabic stop+sonorant, which results from prothesis (*tra* → **ət.ra*), yields bad syllable contact. Predictions 2 and 3 – that prothesis is preferred over anaptyxis for *s*+stop, and that among *s*+sonorant sequences, more anaptyxis is observed as the sonorant increases in sonority – are explained as follows. When C in an sC cluster is low in sonority, good syllable contact results from prothesis (*sta* → *əs.ta*). However, as the sonority of C increases, prothesis results in poor syllable contact (*sla* → **əs.la*); thus, anaptyxis is preferred (*sla* → *əs.la*).

7. Interim summary

To summarize thus far, I have returned to a structural approach to syllable well-formedness, but one which is informed by perceptual considerations. From a perceptual perspective, (i) segments strive to be ordered to maximize their perceptibility, but this can be overridden by structural well-formedness considerations; and (ii) in contrast to other obstruents, *s* (and sibilants more generally) is special because it has strong internal cues, which ensures that it can be accurately identified, independent of the context in which it occurs. From a structural perspective, (word-initial) sC clusters are analysed as coda+onset which, when coupled with the Syllable Contact Law, accounts for the following: (i) no language that permits sC clusters forbids *s*+stop; (ii) sC clusters decrease in well-formedness as the sonority of C increases; and (iii) patterns of epenthesis (prothesis versus anaptyxis) are sensitive to syllable contact and the options available for the analysis of various types of clusters (*s*-initial vs. obstruent-initial).

8. Second language acquisition of sC cluster well-formedness

In the remainder of the paper, I turn to explore the predictions of a coda+onset analysis of sC clusters for second language acquisition. When learners whose native language lacks sC clusters are faced with the challenge of acquiring a superset grammar on this dimension (i.e., the L2 is more permissive than the L1), two predictions are as follows:

---

[7] Unlike what is observed here for French, sC clusters in initial and medial position do not always respect the same constraints in a given language. In English, for example, (11) shows that initial sC with C of all sonority profiles is attested, while word-medially, *s*+sonorant is virtually unattested, similar to French (see Goad 2012). The inverse holds of Spanish: word-initial sC is out altogether, but word-medially, Spanish has the profile that Dutch exhibits word-initially (see (11)). Although initial and medial sC clusters have the same structure, they are preceded by an empty and filled nucleus, respectively, which licenses differences of this sort. Importantly, the typology of medial sC clusters exactly mirrors the profile seen for word-initial sC in (11): *s*+stop is attested in all languages that permit sC and the cluster gets worse as the sonority of C increases, as mentioned in the text.

- learners should acquire *s*+stop before *s*+sonorant in the L2 (partly confirmed; see Goad to appear);
- learners who are exposed only to the well-formedness of *s*+sonorant in the L2 should infer the well-formedness of *s*+stop.

When learners whose native language contains sC clusters are acquiring a subset grammar on this dimension (i.e., the L2 is more restrictive than the L1), one prediction is as follows:

- if learners are exposed only to the ill-formedness of *s*+stop in the L2, they should infer the ill-formedness of *s*+sonorant.

This last prediction is examined here, within the broader context of the types of evidence that may be available to learners acquiring a subset grammar in phonology.

9. Evidence

The types of evidence available indicating the (un)suitability of the L1 grammar that is transferred into the learner's interlanguage grammar will depend on the shapes of the L1 and L2 grammars and, as we will see below, on the acquisition context. Consider two languages that differ in their grammars of left-edge clusters, Brazilian Portuguese (BP) and English. As was observed in (11), BP lacks word-initial sC clusters altogether while English permits sC regardless of the sonority value of the C following *s*.

When learning a superset language, for example when considering the acquisition of sC clusters in English-type languages by L1 speakers of BP-type languages, direct positive evidence is available: discovering that a structure is well formed from the presence of appropriate forms in the ambient data. Assuming that these forms can be accurately perceived, acquisition should ultimately be successful.

When learning a subset grammar, for example when acquiring the ill-formed status of sC clusters in BP-type languages by speakers of English-type languages, direct positive evidence *may* be available (from morphophonemic alternations, loan word adaptations, or a comparison of cognates; see Schwartz & Goad (2015) for discussion), but it isn't always accessible (e.g. the use of alternations requires a solid understanding of the morphological structure of words, which learners may not have). When this type of evidence isn't available or isn't accessible, must learners rely on negative evidence, either indirect negative evidence (inferring that a structure is ill-formed through the absence of relevant forms from the ambient data) or direct negative evidence (explicitly being told that a particular type of form is illicit)? We suggest that the answer is no, that another source of evidence may be available in certain acquisition contexts: indirect positive evidence (IPE). IPE is evidence from errors in the learner's L1 made by native speakers of the learner's L2 (Schwartz & Goad 2015).

We exemplify IPE through the following acquisition context. Consider a native speaker of English who is acquiring BP. It is highly possible that the English speaker may come across a native speaker of BP, who is in the process of acquiring English. If (part of) their conversation takes place in English and the BP speaker is not highly proficient in this language, the BP speaker will likely epenthesize a vowel before some or all sC-initial words (e.g. *sky → isky*). If epenthesis is detected by the English speaker, s/he may infer that the constraints of BP do not allow sC clusters in word-initial position. Because evidence for the constraints that hold of one language, BP, is accessible via errors made in another language, English, this type of evidence is indirect.

10. Experiment
10.1. Goals and predictions

Since IPE is present in accented speech, our goal was to test whether it is potentially an effective source of evidence for second language learners. To do this, we experimentally examined whether English speakers who are naïve to the structure of BP can deduce phonotactic information about BP after being exposed to IPE in the form of BP-accented English.[8] To ensure as well that any judgements that English speakers make result from the construction and application of rule-based generalizations,

---

[8] The experiment reported on here appears in expanded form in Schwartz & Goad (2015).

rather than analogy, a second goal was to test for generalization beyond the data to which learners were exposed and, specifically, to see if this behaviour was consistent with cross-linguistic constraints on the well-formedness of sC clusters (*s*+stop vs. *s*+sonorant).

The experiment was designed to test the following predictions (among others):

1. Participants will use IPE to conclude that word-initial sC clusters are ill-formed in BP;
2. Participants exposed only to the ill-formedness of *s*+stop will conclude that *s*+sonorant is also ill-formed;
3. Participants will not overgeneralize the repair pattern they employ for word-initial sC clusters to obstruent-initial clusters in this position.

10.2. Methodology

32 native speakers of Canadian English (age range 19-33), with no more than low-intermediate proficiency in another language participated in the study. They had no knowledge of Portuguese or any other language with similar repair of word initial clusters (notably Spanish). Participants could not have lived in the United States for an extended period of time, because of the more common exposure to Spanish and Spanish-accented English in the US.

Participants were trained by listening (over headphones) to one of two 7-minute dialogues between two native speakers of BP speaking BP-accented English; they were told that the speakers' L1 was Samoan, to better distract them from possibly identifying the accented speech as something they had heard before.[9]

All sC clusters in the dialogues appeared with prothetic [i] (e.g. *istop*). Each dialogue contained 48 tokens of word-initial sC clusters; the place and manner of articulation of C were controlled. There were, in addition, 30 tokens of obstruent+liquid (CL) clusters in each dialogue, which were not repaired. Indeed, epenthesis was limited to the sC context, as our goal was to test whether IPE is both accessible and interpretable given limited exposure to the language.

Participants were randomly assigned to one of two conditions that differed in the types of sC clusters present in the dialogue. One dialogue contained both *s*+stop and *s*+sonorant initial stimuli (8 tokens each of *sp, st, sk, sm, sn, sl*); this is referred to as the Sonorant condition. The other dialogue contained only *s*+stop initial stimuli (16 tokens each of *sp, st, sk*); this is referred to as the NoSonorant condition. Since all sC clusters were repaired in both dialogues, participants in the Sonorant condition received evidence for the illicit status of sC from anaptyxis applying to both *s*+stop and *s*+sonorant initial words, while those in the NoSonorant condition received evidence for the illicit status of initial sC clusters only from *s*+stop.

After training, participants performed a judgement task. They saw 120 novel words (displayed orthographically on a computer screen) that began with sC, CL, or C and were asked to produce each word with the same accent as heard in the dialogue. All participants judged the same stimuli, regardless of which condition they were assigned to.

10.3. Results

Prediction 1 states that participants will be able to use IPE to determine that word-initial sC clusters are ill-formed in BP. Figure 2 overleaf provides the results: Repair (in black) indicates appropriate responses, cases where sC clusters were repaired, consistent with the data to which participants were exposed; grey indicates inappropriate responses, cases where sC clusters were produced as English-like. As the figure suggests, 11 participants performed significantly higher than

---

[9] At the end of the experiment, participants were told that the native language was not Samoan, and were asked to guess the language, to probe for any connection between their performance on the judgement task (see below) and identification of the native language as Portuguese or Spanish. Several participants who performed well in the judgement task failed to identify the language as Portuguese or Spanish and several who performed poorly did identify the language as Portuguese or Spanish. Thus, we do not think that previous exposure to Portuguese, Spanish or Portuguese/Spanish-accented English significantly impacted the results (see Schwartz & Goad (2015) for further details).

chance (determined by a binomial test) ($ps \leq 0.001$); 19 participants performed significantly lower than chance ($ps \leq 0.002$); and the performance of 2 participants was neither significantly higher ($p = 0.061$) nor significantly lower than chance ($p = 0.088$). Henceforth, we will refer to these three groups of participants as 'learners', 'non-learners' and 'chance performers' respectively. We can conclude that Prediction 1 is supported but only for the learners.[10]
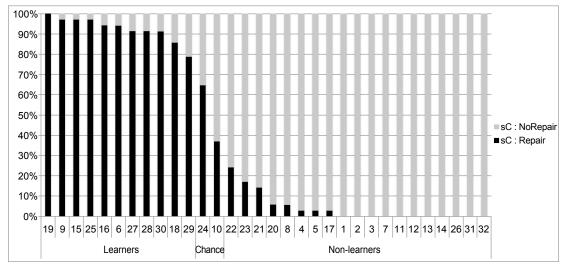
Figure 2. Responses to all sC-initial words.

Prediction 2 states that participants exposed only to the ill-formedness of *s*+stop will conclude that *s*+sonorant is also ill-formed. Recall that participants were assigned to two conditions. The behaviour of those in the NoSonorant condition speaks to Prediction 2.

If the coda+onset analysis of sC clusters is correct, then syllable contact should inform cluster well-formedness. As was observed in (11), if *s*+stop is ill-formed in a language, then *s*+sonorant will be ill-formed as well. Participants in the NoSonorant condition were provided with no information on the well- or ill-formedness of *s*+sonorant; there were simply no words of the appropriate shape in the dialogue to which they were exposed.

Of the 16 participants in the NoSonorant condition, 8 were learners. We thus focus on their results. Figure 3 shows that these 8 learners appropriately generalized the constraint against *s*+stop to *s*+sonorant and repaired the latter clusters as well as the former. A paired t-test reveals that their treatment of the two types of clusters was not significantly different (t = 1.8825, df = 7, $p = 0.1018$). Prediction 2 is thus supported for all 8 learners in the NoSonorant condition.

Prediction 3 states that participants will not overgeneralize the repair pattern they employ for word-initial sC clusters to CL clusters in this position. No overgeneralization is expected for two reasons: one, the data to which participants were exposed indicated that CL is well-formed; two, if learners failed to notice this, on grounds that noticing the absence of something (i.e., CL also does not undergo repair in the participants' L1) is more difficult than noticing the presence of something unexpected (i.e., epenthesis before sC, from the perspective of the participants' L1), no overgeneralization would be expected on a structural account of cluster well-formedness, as different representations hold for CL and sC clusters ((3a) and (10d)). Prediction 4 was supported for those participants who fell into the learner group; there was only one instance of repair across the 11 learners.[11]

---

[10] The learners appropriately repaired the sC clusters using prothesis (e.g., *skesu* → [iskesu]). The repairs of the less proficient participants, both the chance performers and the non-learners, were more variable, involving, in addition to prothesis, anaptyxis ([sikesu]), deletion ([sesu] or [kesu]), and metathesis ([seksu]).

[11] A number of participants did unexpectedly overgeneralize the repair to words that began with singleton *s*. See Schwartz & Goad (2015) for further discussion and explanation for this pattern of behaviour.
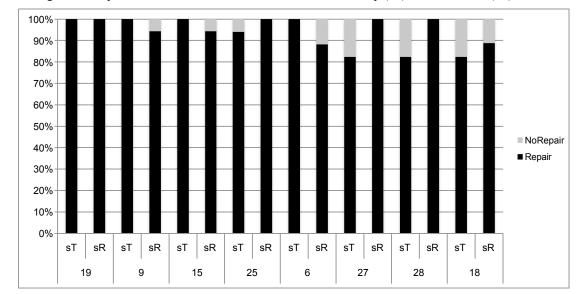
Figure 3. Responses of learners in NoSonorant condition: *s*+stop (sT) vs. *s*+sonorant (sR) words.

11. Interim Summary

To summarize, we have argued that when learning a subset grammar, in some acquisition contexts, learners can use indirect positive evidence: evidence from errors in the learner's L1 made by native speakers of the learner's L2. We have seen that some individuals can use IPE after short exposure to the L2. This is manifested through acquisition of the constraint against sC and use of the appropriate repair (prothesis). All participants who used IPE and who were exposed only to the ill-formedness of *s*+stop appropriately generalized the constraint against sC to *s*+sonorant, respecting cross-linguistic constraints on the well-formedness of sC clusters.

12. Conclusion

In this paper, I have motivated a structural approach to sC clusters. Perception has been shown to play a role in syllable well-formedness, but perceptibility considerations can be trumped by structural considerations. In spite of their position, word-initial sC clusters were analysed as coda+onset, which correctly predicts that their phonotactic behaviour mirrors that of word-internal coda+onset clusters, as determined by the Syllable Contact Law. Specifically, *s*+stop is preferred over *s*+sonorant, while branching onsets have the inverse profile. Further motivation for the phonotactic constraints of sC came from second language acquisition: learners exposed only to the ill-formedness of *s*+stop correctly inferred the ill-formedness of *s*+sonorant, supporting the view that interlanguage grammars are possible grammars in the cluster domain.

References

Clements, George Nick. 1990. The role of the sonority cycle in core syllabification. In John Kingston & Mary E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech.* Cambridge: Cambridge University Press, pp. 283-333.

Côté, Marie-Hélène. 2000. *Consonant cluster phonotactics: A perceptual approach.* Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.

Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics* 7: 71-116.

Goad, Heather. 2011. The representation of sC clusters. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology.* Oxford: Wiley-Blackwell, pp. 898-923.

Goad, Heather. 2012. *s*C clusters are (almost always) coda-initial. *The Linguistic Review* 29: 335-373.

Goad, Heather. to appear. Superset and subset grammars in second language acquisition: The role of sonority in the representation of sC clusters. In Pedro Guijarro-Fuentes & Maria Juan-Garau (eds.), *Acquisition of Romance languages: Old acquisition challenges and new explanations in generative perspective.* Berlin: De Gruyter Mouton.

Goad, Heather & Yvan Rose. 2004. Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition.* Cambridge: Cambridge University Press, pp. 109-157.

Goldsmith, John. 1990. *Autosegmental and metrical phonology.* Oxford: Blackwell.

Gouskova, Maria. 2001. Falling sonority onsets, loanwords, and syllable contact. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *Proceedings from the Main Session of the 37th Meeting of the Chicago Linguistic Society.* Chicago, IL: Chicago Linguistic Society, pp. 175-186.

Hall, Tracy Alan. 1992. *Syllable structure and syllable-related processes in German.* Tübingen: Max Niemeyer Verlag.

Hulst, Harry van der. 1984. *Syllable structure and stress in Dutch.* Dordrecht: Foris.

Karimi, Simin. 1987. Farsi speakers and the initial consonant cluster in English. In Georgette Ioup & Steven Weinberger (eds.), *Interlanguage phonology: The acquisition of a second language sound system*. Cambridge, MA: Newbury House, pp. 305-318.

Kaye, Jonathan. 1992. Do you believe in magic? The story of *s*+C sequences. *SOAS Working Papers in Linguistics and Phonetics* 2: 293-313.

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

Murray, Robert & Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59: 514-28.

Parker, Steve. 2002. *Quantifying the sonority hierarchy.* Amherst, MA: University of Massachusetts doctoral dissertation.

Schwartz, Misha & Heather Goad. 2015. Indirect positive evidence in the acquisition of a subset grammar. Submitted.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In Mark Aronoff & Richard Oehrle (eds.), *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students.* Cambridge, MA: MIT Press, pp. 107-136.

Smith, Jennifer. 2005. *Phonological augmentation in prominent positions.* New York: Routledge.

Steriade, Donca. 1982. *Greek prosodies and the nature of syllabification.* Cambridge, MA: MIT doctoral dissertation.

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In Osamu Fujimura, Brian Joseph & Bohumil Palek (eds.), *Proceedings of the 1998 Linguistics and Phonetics Conference*. Prague: Karolinum, pp. 205-242.

Vaux, Bert & Andrew Wolfe. 2009. The appendix. In Eric Raimy & Charles Cairns (eds.), *Contemporary views on architecture and representations in phonology*. Cambridge, MA: MIT Press, pp. 101-143.

Wright, Richard. 1996. *Consonant clusters and cue preservation in Tsou.* Los Angeles, CA: University of California Los Angeles doctoral dissertation.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*. Cambridge: CUP, pp. 34-57.

Proceedings of the 33rd West Coast Conference on Formal Linguistics

edited by
Kyeong-min Kim, Pocholo Umbal,
Trevor Block, Queenie Chan,
Tanie Cheng, Kelli Finney, Mara Katz,
Sophie Nickel-Thompson, and Lisa Shorten

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Goad, Heather. 2016. Phonotactic Evidence from Typology and Acquisition for a Coda+Onset Analysis of Initial sC Clusters. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong-min Kim et al., 17-28. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3221.