UiT The Arctic University of Norway

Faculty of Humanities, Social Science and Education

The acquisition of complex onsets in Icelandic: the effects of markedness, sonority, and frequency

A corpus-based study

Jagoda Anna Dyga

Master’s thesis in Theoretical Linguistics, LIN-3990, May 2023

Acknowledgments

This thesis would not have been possible without the many wonderful and kind people who supported me during the whole year of working on this project. Foremost, I would like to express my gratitude to my supervisor, prof. Martin Krämer. I want to thank him for all his invaluable guidance, patience, support, and encouragement at every stage of the whole project. I would like to thank my co-supervisor, dr. Olga Urek for her invaluable feedback and guidance. I am also grateful to my second co-supervisor, prof. Björn Lundquist for his insightful comments and suggestions about statistical analysis of my data, as well as other parts of my thesis.

I want to thank prof. Sergey Minor, from whom I also received helpful feedback and numerous comments on the statistical analysis. I also want to thank Andrea Harrison, my friend and colleague, for taking the time to proof-read my thesis. I would like to thank everyone for their time, encouragement, and genuine interest in my research. It was a great pleasure to be part of the linguistic community at the University of Tromsø, as well as the SALT project. To have the opportunity to meet so many inspiring people willing to share their passion for linguistics and research has motivated me to persevere.

Also, I would like to thank prof. Þóra Másdóttir for creating the Másdóttir corpus. She has made an extensive (and exciting) dataset available for further research, and her personal correspondence has been of great help to me. Without the Másdóttir corpus, this project could not have been accomplished.

I would like to acknowledge fellow students for their great company during the many long hours spent at the office while working on our projects. Also, I would like to thank my family and friends for always being there for me when I needed them. I want to thank my parents for their love and support.

Lastly, I want to thank my partner Zebastian, for whose presence in my life I am deeply grateful. Thank you for always believing in me, being patient with me even in the most difficult moments, and showing your love and care in all the small and big gestures.

Abstract

The acquisition of consonant clusters is determined by several underlying factors, and recent studies demonstrated that input frequency, investigated in isolation, cannot account for the order of acquisition (e.g., Jarosz 2017). The study conducted in this thesis was an empirical test of the theories of phonological acquisition. It investigated the simplification patterns occurring in initial clusters and the role of sonority, type frequency, and articulatory complexity of individual segments on the order of their acquisition. To investigate the influence of these factors on accuracy in producing initial consonant clusters, I analyzed data from 68 monolingual children (aged 2;6-4;3) acquiring Icelandic, available as a part of the Icelandic Másdóttir corpus (Másdóttir 2014, Másdóttir et al. 2021), which is a part of the PhonBank corpora (Rose & MacWhinney 2014). The investigation of simplification patterns revealed a discrepancy in processes targeting branching onsets and sC clusters. Additionally, the results showed that Icelandic children produced clusters composed of stop and liquid as two segments most frequently. Fricative-initial clusters were most frequently realized as one consonant. I demonstrated how two sonority-based generalizations, the Minimal Sonority Distance and the Sonority Dispersion Principle, did not account for all the tendencies in the order of acquisition, and could not predict children's low accuracy on fricative-initial clusters. The type frequencies of initial onset clusters in Icelandic did not correlate with children's accuracy. The findings revealed that what has previously been attributed to the role of sonority, could be explained on the basis of articulatory complexity of individual cluster members, or possibly perceptual cues.

Table of Contents

List of Tables

List of Figures

1. Introduction

Over the recent years, the acquisition of consonant clusters has been of interest not only to phonologists and linguists investigating the acquisition of a first language, but also to speech pathologists. Uttering a cluster of multiple consonants requires advanced motor control, and therefore, consonant clusters are often stated to be one of the most challenging structures to acquire (Smit et al. 1990, McLeod et al. 2001, McLeod & Crowe 2018). Hence, the accurate production of consonant clusters has often been a target in the treatment of speech disorders (Dodd & Iacono 1989). This necessitates research on cluster acquisition that focuses on investigating the underlying factors in the acquisition process (e.g., Jarosz 2017).

The theories that investigate acquisition of a phonological system of a first language aim to explain how children achieve adult-like competence, and to show that the grammar of each language can be learned by a child (Kager 1999, Smith 2010). At the initial stage of acquiring grammar, early words produced by children strongly differ from the target forms by having a simplified structure, which was first observed by Jakobson (1941/1968). One type of segment that is frequently re-shaped is consonant clusters, which are usually simplified to one consonant. Hence, the first syllables produced by children have a simple unmarked structure and are constructed of a single consonant and a vowel (CV) (Levelt et al. 2000). Gradually within the course of acquisition, children acquire more complex structures, and in the end stage they achieve adult-like proficiency. Since the introduction of Optimality Theory by Prince & Smolensky (1993/2004), research on language acquisition in the last two decades has focused on explaining the process of acquiring the grammar in terms of constraint re-ranking (Demuth, 2011). It is assumed that at the initial stage, markedness constraints outrank faithfulness constraints responsible for faithful mapping between input and output (Demuth 1995, 2011; Gnanadesikan 2004). Thus, it has been proposed that cluster simplification is caused by the influence of constraints which reinforce less marked forms in children's early production (Pater & Barlow 2003, Goad & Rose 2004). Thus, in Optimality Theory, child and adult grammar can be analyzed with the same set of constraints. These constraints represent universal principles that are argued to affect both developing and mature grammar (Pinker 1984, Gnanadesikan 2004).

The aim of this thesis is to contribute to the research on phonological acquisition, and to investigate which factors affect the process of acquiring initial consonant clusters. The notion of sonority is a cornerstone of the Sonority Sequencing Principle, the Minimal Sonority

Distance, and the Sonority Dispersion Principle. These generalizations have been argued to shape the organization of segments within a syllable in languages, and that they affect the course of acquiring phonological system. The findings of many studies demonstrated that sonority plays a role in determining the order of acquiring consonant clusters and cluster simplification processes occurring in early speech (e.g., Fikkert 1994, Gnanadesikan 2004 Barlow 2016, Jarosz 2017). Nevertheless, other studies revealed that several other factors may shape phonological development, such as lexical frequency and the articulatory difficulty of specific segments (e.g., Jarosz et al. 2017, Kistanova 2021). The study conducted in this thesis will test the predictions about the acquisition of initial onset clusters made by the sonority-based generalizations, the frequency effect, and articulatory difficulty by means of statistical modeling. To test all the predictions on a large data sample, I investigate the data from children acquiring Icelandic available as a part of the Másdóttir Corpus (Másdóttir 2014, Másdóttir et al. 2021). I address the following research questions:

1. Is there a difference between reduction patterns occurring in branching onsets and sC clusters in Icelandic?
2. Do the Sonority Dispersion Principle and the Minimal Sonority Distance make the correct predictions regarding the order of acquisition of initial onset clusters in Icelandic?
3. Do the type frequencies of initial consonant clusters in Icelandic affect children's accuracy in producing initial onset clusters?
4. Did children who acquired more singleton consonants also acquire more initial onset clusters?
5. Is there a correlation between the order of acquisition of natural classes and the order of acquisition of initial onset clusters within these classes?

The thesis is organized into seven main chapters. Chapter 2 presents the theoretical background. Before turning to sonority and sonority-based generalizations, I briefly discuss the notion of syllable – a necessary preliminary for discussing phonotactics principles. Further, I discuss two theoretical principles: the Sonority Dispersion Principle and the Minimal Sonority Distance, the predictions of which are investigated in the current study. Next, I discuss the course of acquiring consonant clusters, then summarize the most important assumptions and theories of phonological acquisition relevant to the current study. At the end of the chapter with theoretical background, I present the scope of the current study in light of the discussed theories, and summarize the relevant aspects of Icelandic phonology, such as consonant inventory and phonotactics. In chapter 3, I present the research questions before describing the methodology and analyzed dataset. In chapter 4, I present the results and the quantitative analysis. Lastly, in chapters 5 and 6, I present the discussion and the conclusion, respectively.

2. Theoretical background

2.1 Syllable phonotactics and Sonority Sequencing Principle

2.1.1 Syllable structure

The syllable is a constituent that serves as a basic unit for grouping and organizing speech sounds within words. It is a prosodic unit, which is essential to phonological theory since it allows one to account for constraints on segment distribution (Kahn 1976; Selkirk 1982; Zec 2007). In other words, the syllable is the domain of phonotactics, which organizes and restricts the occurrence of speech sounds in the given positions (Zec 2007). To illustrate, the monosyllabic word *skarp* is a possible word in Norwegian, while **ksapr* and **rpask* cannot be possible words. Both illicitly formed words consist of sequences of consonants which are not allowed on the left and right edge of the syllable in the Norwegian language. In the same fashion, *print* is a possible English word, while **rpitn* is not. Thus, the notion of syllables is necessary to account for the distribution of speech sounds in natural human languages (henceforth languages).

The syllable's internal structure is essential to capture generalizations on segment distribution. It is divided into the onset and the rhyme, composed of the nucleus and the coda. The nucleus is therefore “the central” part of every syllable, and the onset and coda are optional syllable margins. Codas and onsets are usually represented as the letter C, corresponding to a consonant, while the nucleus is represented as the letter V, corresponding to a vowel (Trubetzkoy 1939/1971, Zec 2007). However, it must be noted that the nucleus is not only restricted to vowels; in some languages, consonants can occur in this position (Zec 2007). The syllable is usually represented as a sigma, as shown in figure (1).

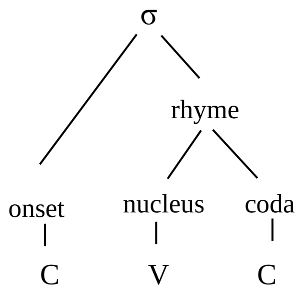


Figure 1. The syllable structure.

The onset, coda and rhyme are building blocks for the syllable. However, not every syllable must contain all of those elements. Clements & Keyser (1983, p.28) listed four main syllable shapes in which the nucleus is the mandatory element: CVC, CV, VC, and V. Languages differ in their syllable inventories; certain languages only allow syllables which have an onset and nucleus, i.e., CV, while other languages have syllables with complex onsets and codas, i.e., (C)(C)V(C)(C) (brackets enclose optional segments). The CV syllable is the most common and the least marked syllable shape, since all languages have this syllable type in their inventory (Clements & Keyser 1983; Zec 2007). The concept of *markedness* refers to the distinction between *marked* and *unmarked* structures, in which the former appears to be simpler, more natural, and more common than the latter (Rice 2007). A markedness relationship between different segments and structures is established based on implicational universals: if one structure is present in a language inventory, then the other, less marked one, will also be present (Carlisle 2006). Vennemann (1988), in *Preference Laws for Syllable Structure*, proposed that the CV syllable is the universal syllable shape. He demonstrated this using historical data, which showed that diachronic change in syllable structure tends towards simplification by causing more complex and, therefore, more marked syllable shapes, to evolve into the unmarked CV shape.

Most languages allow one consonant before and after the nucleus. However, codas and onsets can be *complex* and consist of multiple consonants e.g., CCV, CCCVC, CVCC (Zec 2007). Evidence from diachronic change and phonological processes shows that languages have a preference for shorter syllable margins. Thus, those processes involve reducing complex onsets and codas into simple ones, rather than increasing their complexity (Blevins 1995, Zec 2007). Moreover, based on comparative studies of languages, Greenberg (1978) exhibited that if a given language has onsets and codas of the length *n* in the inventory, it also has onsets and codas if the length *n-1.* That is, no language has only the syllable type CCVCC in its inventory, but no syllables of the type CVC (Carlisle 1998). This generalization has only one exception: the presence of V in the language inventory does not imply the presence of CV (Greenberg 1978). Thus, complex syllable margins are stated to be less preferred and more marked.

The domain of syllables is crucial for stating language-specific phonotactics. As demonstrated in the first paragraph, languages allow only certain segments to co-occur in the coda position, while other sequences of segments can form a complex onset. Grouping segments into syllables is mainly governed by sonority, which is described in the next subsection.

2.1.2 Sonority and Sonority Sequencing Principle

The sonority of speech sounds plays a role in shaping syllables across languages, which has been widely observed and discussed in the literature (e.g., Steriade 1982, Selkirk 1984, Clements 1990). Segments which exhibit similar properties cross-linguistically are grouped by sonority, which can be defined *"…as a unique type of relative, n-ary (non-binary) feature-like phonological element that potentially categorizes all speech sounds into a hierarchical scale*" (Parker 2011, p. 1). Sonority has a special status among phonological features, since every speech sound bears an inherent value for sonority, and unlike other features, this value cannot spread (Parker 2011). It is a phonological property but has been shown that sonority also has cross-linguistic physical, and acoustic correlates. It corresponds to the resonance of sounds, and it can be measured by acoustic energy, loudness, and intensity (Parker 2008, Clements 2009). Moreover, the sonority level corresponds to the degree of obstruction of the airstream occurring during the articulation of sounds. Thus, the classes of sounds, which are grouped by the level of sonority, can be divided into phonetic categories of the manner of articulation (Carr 1993, Parker 2011).

Since all phonemes are assigned a sonority value, speech sounds can be organized into a sonority scale. The Sonority Hierarchy states that vowels are segments of the highest sonority, followed by glides, liquids, and nasals (Sievers 1881, Saussure 1916, Clements 1990). Obstruents are at the bottom of the scale, considered to be the segments of the lowest sonority. Within this class, it is claimed that fricatives have a higher sonority value than stops (Steriade 1982). The version of the Sonority Hierarchy that has been adopted by most scholars, and will be adopted for the current study, is presented in figure (2) (based on Steriade 1982, Clements 1990). The numeric value given in the bracket is known as *a sonority index* (Parker, 2002).

Most sonorous ↑ (6) vowels
(5) glides
(4) liquids
(3) nasals
↓ (2) fricatives
Least sonorous (1) stops

Figure 2. The Sonority Hierarchy

Nevertheless, there are more detailed proposals for the scale in the literature. For instance, Kiparsky (1979) divides liquids into lateral and rhotic sounds, while in the scale proposed by Selkirk (1984), there is a distinction between voiced and voiceless segments. Parker (2002), in his dissertation, summarizes different proposals of the sonority scale and raises the question of whether the Sonority Hierarchy should be considered to be universal, or rather specific to each language. Based on phonetic and phonological evidence, he distinguished 17 sonority levels and demonstrated that a universal Sonority Hierarchy could be established. However, it must be noted that more recent work suggests that the sonority scale might have a language-specific distribution. Krämer & Zec (2020), based on the analysis of over 200 languages, show evidence for languages having either one or two types of nasals, which differ in the degree of sonority, active in their phonological systems. Thus, nasals might occupy two different places on the sonority scale.

In the organization of syllable structure, sonority governs which segments occupy margins and nuclei within syllables. A well-formed syllable must include a sonority rise from the onset to the nucleus, and a sonority fall from the nucleus to the coda, as stated by *the Sonority Sequencing Principle* (SSP) (Selkirk 1984, Clements 1990). Furthermore, it is claimed that sonority shapes the organization of segments within complex onsets and codas: segments of the lowest sonority occupy the margins. In contrast, segments with higher sonority level occupy the position closer to the syllable peak (Jespersen 1904). Figure (3) demonstrates an example of a syllable as a sonority rise (based on figure found in Clements 1990, p. 299).

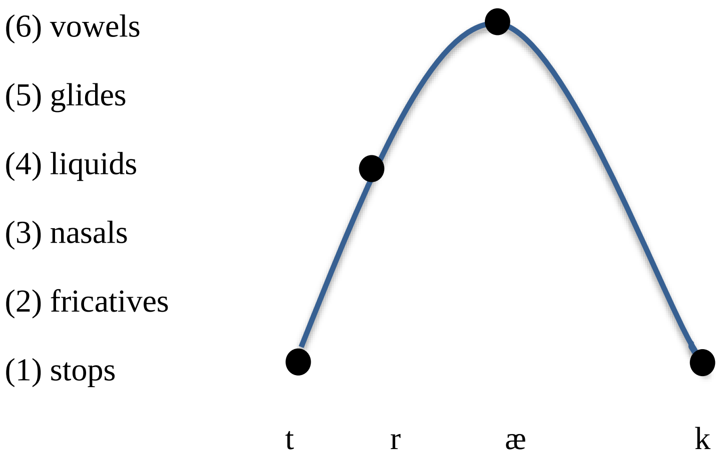


*Figure 3.*The syllable as a sonority peak.

Consonant clusters following the SSP can be found in languages with tautosyllabic clusters. Many languages only include biconsonantal clusters obeying the SSP in their inventories (Clements 1990; Parker 2012, p. 103). These typological observations lead to the proposal that SSP-adhering clusters are the least marked and are universally preferred. Furthermore, clusters with sonority plateaus are stated to be less marked than sonority reversals (Parker 2011, 2012). In an empirical study conducted by Frisch (2015), the SSP generalization was shown to account for 72% of consonant clusters in 47 languages. Moreover, it has been shown that the SSP does not only shape syllable phonotactics cross-linguistically, but also is active in the speaker's grammar while dealing with unattested consonant clusters. Several studies demonstrated that speakers of languages with reduced cluster inventories prefer clusters that obey the SSP (e.g., Berent et al. 2012, Zhao & Berent 2016). Therefore, sonority projection effects accounts for sonority, or the SSP being encoded in Universal Grammar, the theory that humans have innate linguistic knowledge (1968) (Parker 2011).

In the next subsections, I discuss two principles that exhaustively determine the composition of tautosyllabic clusters beyond the coarse SSP. These two phonological generalizations regarding sonority are central to the study conducted in this thesis, and their assumptions are tested on children's data.

2.1.3 The Sonority Dispersion Principle and The Minimal Sonority Distance

The SSP generalization establishes a markedness relationship between sonority rises, falls, and plateaus. However, it does not state explicitly which sonority profiles within each of these classes are more complex than others. This section will introduce two sonority-based generalizations that account for relative complexity of consonant clusters: the Sonority Dispersion Principle and The Minimal Sonority Distance. Both generalizations are based on implicational universals of consonant clusters in languages. Since the markedness relation of onset clusters is the most essential for the study presented in this thesis, I only discuss these generalizations as far as they pertain to complex onsets.

To account for the relative complexity of syllables with sonority rises, Clements (1990) introduced the Dispersion Principle, which states that the optimal syllable has the maximal, and the most evenly dispersed, rise in sonority. In this generalization, Clements referred to the notion of a *demisyllable.* A syllable can be divided into two demisyllables which overlap, with the nucleus being the shared element of the initial and the final demisyllable. To illustrate, the word *track* /træk/ can be divided into two demisyllables – /træ/, and /æk/. This notion allows to account for sonority profiles in onsets and codas independently from one another. The dispersion value presented in Clements (1990) was calculated based on the sonority hierarchy given in (1), which does not distinguish between two classes of obstruents.

(1) Obstruents (O) < nasals (N) < liquids (L) < glides (G) < vowels (V)

The dispersion value (D) of demisyllables is calculated based on the formula given in (2), where *m* is equal to n(n-1), n is equal to the number of segments, and *d* is equal to the sonority distance between each pair of segments (Clements 1990, p. 304).

(2) Sonority Dispersion formula

$$D = \sum_{m}^{i-1} \frac{1}{d_i^2}$$

The dispersion value rises together with the number of segments which constitutes a complex onset, and a bigger value indicates the increase in the relative complexity of a syllable margin. Therefore, demisyllables of the CCCV shape, thus, consisting of four segments have higher D value than three-segment demisyllables in the form CCV. The dispersion values, calculated assuming the sonority hierarchy given in (2), are presented in table (1) which gives the respective complexity ranking of initial onset clusters (based on Clements 1990, p. 305-307).

Complexity ranking	Sonority profile	Dispersion Value
1	*OLV*	*0,56*
2	*ONV, OGV*	*1,17*
3	*NLV, NGV*	*1,36*
4	*LGV*	*2,25*
5	*ONGV*	*2,53*
6	*OLGV, ONLV*	*2,67*
7	*NLGV*	*3,61*

Table 1. The Sonority Dispersion value of initial onset clusters

Thus, according to the Sonority Dispersion Principle, the most preferred demisyllable is composed of an obstruent, a liquid, and a vowel in the nucleus position (OLV). On the scale given in (1), liquids lie exactly between obstruents and vowels, so clusters with the OLV profile have the most equally dispersed sonority. It is important to note that Sonority Dispersion does not account for sonority plateaus and reversals, which are commonly attested in languages.

The second generalization, which establishes markedness relationships within complex syllable margins, is known as the Minimal Sonority Distance (Steriade 1982, Selkirk 1984, Zec 2007). This assumption accounts for the fact that languages often only allow consonant clusters, which contain consonants separated by a certain number of ranks on the sonority scale. To illustrate, clusters with big sonority rises, such as OL and OG (e.g., /pl/, /bj/), are more common than clusters with a relatively smaller sonority rise, such as NL, or LG (e.g., /mr/, /rw/) (Parker 2011). The Minimal Sonority distance is stated in (3) (adopted from Parker 2011, p. 9).

(3) Minimal Sonority Distance (MSD)
Given an onset composed of two segments, C_1 *and* C_2*, if a = Sonority Index of* C_1 *and* $b = SI(C_2)$*, then* $b - a \geq x$*, where* $x \in \{0, 1, 2, 3\}$.

Thus, based on the MSD generalization, the most preferred complex onsets are those comprising the biggest sonority rise of 3 sonority ranks. According to the typology based on the MSD, if a language allows tautosyllabic clusters with a smaller sonority rise, it will also allow clusters with a respectively bigger sonority distance (Parker 2011, 2012) Table (2) presents the markedness relationship of biconsonantal onset clusters according to the MSD based on the sonority scale stated in (1) (based on Parker 2011, p. 9).

Complexity ranking	Sonority profile	A rise in sonority
1	*OG*	*MSD=3*
2	*OL, NG*	*MSD=2*
3	*ON, NL*	*MSD=1*
4	*OO, NN, LL, GG*	*MSD =0*

Table 2. The Minimal Sonority Distance of initial onset clusters.

Contrary to the Sonority Dispersion Principle, the MSD generalization does not address the sonority index of the following vowel. It is because language phonotactic restrictions usually do not target the sonority distance between the offset consonant of the cluster and the nucleus

(Blevins 1995). The main difference between these two formal devices lies in determining the least marked complex onset: the Sonority Dispersion Principle favors clusters with the sonority profile OL. In contrast, the MSD approach favors OG clusters. Thus, the question arises: which approach makes the correct predictions regarding universally preferred complex onsets? Parker (2012) addressed this issue and conducted a typological survey investigating the inventories of 122 languages with complex onsets. The results exhibited that both generalizations can find validation in the empirical data. However, glides were more likely to occur as a second member of a complex onset, rather than liquids. Parker's (2012) study did not reveal whether this observation is statistically significant. Thus, the question of which sonority profile of complex onset should be considered as universally preferred remains unsolved. In the study presented in this thesis, I investigate whether the Sonority Dispersion Principle and the Minimal Sonority Distance will be supported by the processes found in the acquisition data.

2.1.4 The status of sC clusters

As mentioned in the previous subsection, not all consonant clusters in languages obey the SSP generalization. In this subsection, I discuss such a cluster type, consisting of a voiceless sibilant /s/ followed by a consonant. It frequently escapes the assumptions of SSP, and it is present in the inventories of many languages (henceforth sC clusters).

A sonority reversal is created when /s/ is followed by an obstruent. When investigating the phonotactic behavior of consonant clusters in languages, it can be observed that the *s + obstruent* sequence is often the only cluster type which violates the SSP (Weijer 1996, Goad 2011). This is the case for English, Icelandic, and several other Germanic languages. Although not all sC clusters violate SSP since *s + sonorant* clusters constitute a sonority rise, some other properties make this cluster type peculiar. In English, homorganic clusters are prohibited and clusters such as /pm/ or /tl/ do not exist in its inventory. However, consonants comprising sC clusters can have the same place of articulation, and clusters such as /sn/, /sl/, or /st/ are well-formed English onsets (Yavaş 2006). Additionally, /s/ is the only consonant that a nasal can follow (Barlow 2001).

Considering these discrepancies in phonotactic behavior, it has been proposed that sC clusters differ in the internal structure from other complex onsets composed of an *obstruent + sonorant*, such as /pl/, /kr/ (henceforth branching onsets). Goad (2011) summarized arguments and proposals which suggest that contrary to branching onsets, sC clusters are right-headed. Furthermore, it has been suggested that /s/ is structurally outside the syllable constituent.

Therefore, it should be analyzed as an appendix or an adjunct, rather than a member of a complex onset (e.g., Steriade 1982, Goldsmith 1990). In this account, there is no violation of SSP since *s* is syllabified outside the syllable onset, as presented in the figure (4) (based on Barlow 2001, p. 11). Alternatively, Kaye (1992), based on the observations from syllabification rules, proposed that /s/ is a coda of an empty syllable.

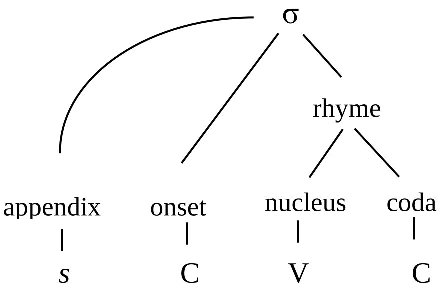


Figure 4. An appendix initial syllable

Evidence from typologies also supports the structural difference of sC clusters. Some languages only have branching onsets but prohibit sC clusters, e.g., Spanish, while in other languages sC clusters are the only allowed complex onsets, e.g., Acoma spoken in New Mexico (Harris 1983, Miller 1965, as cited in Goad 2011). Vaux & Wolfe (2009) demonstrated that evidence for extrasyllabicity comes not only from typological observations and phonological processes, but also from domains such as psycholinguistic experiments or speech errors. Experimental studies showed that native speakers tend to syllabify sC clusters as non-tautosyllabic, unlike branching onsets (e.g., Treiman & Zukowski 1990). Another piece of evidence comes from speakers with an aphasic speech disorder who tend to reduce clusters. A study conducted by Romani & Calabrese (1998) showed that *s* + *stop* clusters are handled significantly differently than branching onsets. These are just a few studies that imply a special status for sC clusters.

As mentioned above, not all sC clusters violate SSP generalization: clusters composed of *s* + *sonorant,* such as /sn/, /sl/, /sr/, have a rising sonority profile. Therefore, it has been proposed that *s* + *sonorant* clusters pattern with other branching onsets (e.g., Hall 1992). According to this proposal, only *s* + *stop* clusters should be seen as structurally different (Goad 2011). This statement is mainly supported by cluster simplification patterns found in developing grammar, which I discuss in detail in the next section. The second piece of evidence is the behavior of *s* + *stop* while being targeted by different phonological processes. Weijer (1996) argued that the sequence of *s* + *stop* patterns differ from other consonant clusters in reduplication patterns, or in loan word phonology. Thus, he proposed that *s* + *stop* sequences should be analyzed as complex segments, analogous to affricates. On the other hand, Steriade (1982) demonstrated that in Attic Greek all sC clusters share phonotactic properties. Thus, it is not apparent whether all sC clusters should be seen as non-tautosyllabic, or rather whether only those comprising a sonority fall are structurally different from branching onsets.

This subsection summarized evidence from the literature that suggests that sC clusters and branching onsets differ in structural properties. Putting aside the disagreement on how those differences should be formally represented, the necessity of a distinct analysis for each of those two cluster types is still striking. In the next sections, I discuss evidence from child language, which forms the basis of the study conducted in this thesis.

2.2 Acquisition of consonant clusters

In this section, the focus is shifted to the process of acquisition of syllable structure and consonant clusters. In the latter part, I discuss general tendencies in cluster acquisition found cross-linguistically. In this section, I focus on the role of sonority in the acquisition of complex onsets, simplification strategies that occur in clusters, and factors impacting the process of acquisition. Crucially, I address the difference between acquiring branching onsets and sC clusters.

2.2.1 The order of acquisition

Consonant clusters are acquired at the last stage of phonological development and are often one of the most difficult structures to acquire, as reported by many speech pathologists (McLeod et al. 2001). Their research shows that even children at the age of 7 struggle to produce clusters correctly (Smit et al. 1990). It is widely documented that at the early stage of speech development, children produce simple, open syllables of the shape CV, and the first words produced by children are usually shaped as CVC, CVCV. (e.g., Stoel-Gammon 1987, Fikkert 1994). As the process of acquisition progresses, children start to produce syllables with more segments, with the first attempts to produce syllables with clusters around two years of age in many languages (McLeod et al. 2001). More complex syllable shapes, such as syllables with triconsonantal onsets, are usually acquired at the last stage after biconsonantal onsets are mastered (Shatz 2019). Uttering clusters requires a lot of motor control and precision of gestural coordination; thus, posing a particular articulatory challenge for children with speech impairment (e.g., McLeod et al. 1997, Dodd & Iacono 1989).

Regarding the order of acquisition of clusters in different languages, it has been observed that sonority-based generalizations can predict which clusters are the first to be mastered. Several researchers addressed the predictions of SSP directly and noted that children acquiring their first language produce clusters with a greater sonority distance more accurately (e.g., Stites et al. 2004 for codas in English, Jarosz 2017 for Polish onsets clusters). Predictions of segment complexity based on sonority have also been applied in speech disorder treatments (Storkel 2018). A study conducted by Gierut (1999) showed that treating the most marked segments predicted by the Minimal Sonority Distance, i.e., clusters with a small sonority distance, ensured greater effectiveness than treating less marked clusters.

On the other hand, a longitudinal case study of a bilingual Russian American girl did not confirm the SSP effect on acquisition order (Kistanova 2021). At the age of two years, she demonstrated a high level of accuracy in production of clusters with sonority reversals and plateaus, but not clusters with a high sonority distance between consonants. Kistanova's (2021) interpretation of these results suggested that other factors, such as the frequency effect and the articulatory difficulty of segments have a greater impact on the accuracy of cluster production than the sonority effect. It has been noted that producing frication, which is necessary to utter fricatives, requires precise control of the vocal tract. Because of their articulatory difficulty, children usually acquire fricatives later than stops and nasals (Kent 1992). McLeod et al. (2011) listed several studies investigating speech development that reported that *stop + liquid* clusters are mastered earlier than *fricative + liquid* clusters. As Stoel-Gammon & Sosa (2007, p. 238) emphasized: *"... phonological acquisition has two basic components: a cognitive–linguistic component associated with learning the phonological system of the ambient language and the development of speech–motor skills needed for adult-like productions".* On the one hand, one could argue that children acquire *stop + liquid* clusters first because it comprises a greater sonority rise than *fricative + liquid*. On the other hand, the discrepancy in the acquisition order reported by McLeod et al. (2011) could also be explained in terms of the phonetic difficulty of fricatives caused by the immaturity of the vocal tract.

2.2.2 Simplification patterns

In their first attempts, children usually fail to produce consonant clusters correctly. Segments in clusters produced at the early stage of phonological acquisition often strongly differ from the target, adult-like forms. McLeod et al. (2001) listed the most commonly occurring phonological processes found in children's speech which occur to simplify clusters. It has been documented in several different languages that these processes usually take place over a period of time, until children master producing target clusters. The most common simplification process is known as *cluster reduction*: the deletion of one or more segments from a cluster. As a result, the cluster is realized as a singleton consonant onset (Grunwell 1987). Cluster reduction has been the most frequently reported process in the acquisition of languages with clusters in their inventories (McLeod et el. 2001). The second typical process can be described as *cluster simplification.* It occurs when children produce both segments of a cluster, but one or more consonants differ from the target pronunciation. Patterns of segment substitution are often driven by processes which also target singleton phonemes in child phonology, such as gliding, stopping, and fronting (Smit 1993). As listed by McLeod et al. (2001), less common processes targeting clusters in child speech, are *metathesis, coalescence*, and *epenthesis*. Metathesis causes the order of segments to be swapped. Coalescence is another type of cluster reduction, and when it occurs, segments are merged into one phoneme that shares features of the original two segments. Epenthesis is the insertion of a vowel, which splits two consonants within a cluster (Johnson & Reimers 2010). More recent studies show that the frequency in which reduction patterns occur can differ cross-linguistically. For instance, the study conducted by Garmann et al. (2020) showed that epenthesis is a typical simplification pattern in child Norwegian.

It has been observed that sonority constraints also affect cluster reduction and simplification. Cluster reduction driven by sonority is a common process found in child phonology (Johnson & Reimers 2010). When clusters are reduced to a singleton consonant, it is usually the least sonorous consonant that is retained. This pattern of cluster reduction was reported by several studies, which argue that it is caused by universal bias for syllables with bigger sonority rises (e.g., Pater 1997, Ohala 1999, Gnanadesikan 2004, Barlow 2016). The deletion pattern results in more preferred syllables, with a bigger sonority rise from the onset to the nucleus, as presented in (4).

(4) Reduction to the less sonorous consonant.
   a. English (from Gnanadesikan 2004, p.77)
      /kliːn/ → [kin] ‘clean’
      /pliz/ → [piz] ‘please’
   b. Polish (from Jarosz 2017, p. 280)
      /znalazw/ → [zala] ‘he found’
      /gwuxa/ → [guka] ‘deaf’

c. German (from Lleó & Prinz 1996, p. 38)
/brɪlə/ → [bel:ə] ‘glasses’
/klainə/ → [kajnə] ‘small’

Regarding cluster simplification, segment substitution often creates a greater sonority distance between two consonants than in the attempted cluster. As mentioned above, gliding is a common process in child language, where liquids are substituted by glides (McLeod et al. 2001). As a result, the produced syllable comprises a greater sonority rise, as presented in (5).

(5) Cluster substitution resulting in a greater sonority rise.
a. Polish (from Jarosz 2017, p. 280).
/mɲɛ/ → [mjɛ] ‘me’
/dva/ → [dwa] ‘two’
b. English (from McLeod et al. 2001, p. 102)
/grin/ → [gwin] ‘green’

Thus, if predictions regarding the impact of sonority on cluster acquisition are correct, the mechanism underlying shaping syllable structure in languages will also influence phonological development. However, even though the sonority pattern of reduction has been well-documented cross-linguistically, there is also evidence that sonority does not influence the speech development of all children. In the study by Lleó & Prinz (1996), Spanish children frequently reduced complex onsets to the more sonorous consonant. These results suggest that despite a strong tendency for sonority-based reduction patterns, this factor alone cannot account for all reduction patterns found in developing grammar. As mentioned earlier, phonetic complexity in segment production might also impact cluster simplification processes. In the cross-linguistic investigation, McLeod & Crowe (2018) analyzed the acquisition order of consonants reported by 64 studies on 27 languages. The result showed that together with fricatives and trills, liquids are acquired at a later stage than stops, nasals and glides. Thereby, if a segment is not acquired yet, it will simply be missing from the target cluster. However, this explanation does not apply to instances where a segment has already been acquired but is still missing from the cluster (Lleó & Prinz 1996). Furthermore, as presented data from Polish show (4a, 5b), liquids and fricatives are not the only segments which are targeted by reduction and simplification; nasals, which are one of the first consonants to be developed, also get reduced to the less sonorous consonant and are substituted with glides. Additionally, the study of McLeod & Crowe (2018) demonstrated that there is a certain degree of variation between individuals. Davis & Bedore (2013) proposed that these discrepancies in the acquisition process can be explained by the interaction of linguistic factors with *child-internal* physical capacities, such as production, perception, and cognition, as well as sociocultural factors. In their view, all these factors must be considered while investigating phonological development.

All afore-mentioned evidence considered, it seems like all of the factors influence the course of acquiring clusters to some extent. It is also not clear which underlying mechanisms have the most significant impact. Furthermore, although general tendencies can be captured, each child will follow a distinct acquisition path due to individual differences and different rates of development.

2.2.3 Acquisition of sC clusters

Unlike the acquisition of branching onsets, the acquisition of sC clusters seems not always to be constrained by the principles of sonority generalizations. Regarding the order of acquisition, several studies reported that children master branching onsets before sC clusters. Fikkert & Freitas (2004) reported that Dutch children acquired *plosive + liquid* and *fricative + liquid* clusters before sC clusters. Also, Smit et al. (1990) investigated the acquisition of clusters by a big group of children, and the results showed that sC clusters are the last ones to be acquired. Children mastered them around the age of 7;00-9;00. This finding might not be surprising, considering that uttering /s/ demands a high degree of motor control, as it requires tongue-tip constriction and laryngeal abduction for devoicing (Koenig et al. 2008). However, some studies reported that sC clusters are mastered before complex onsets. For example, Fikkert & Freitas (2004) compared the cluster development in Dutch and European Portuguese, languages with similar inventories of onset clusters. Children acquiring European Portuguese mastered sC clusters earlier than Dutch children. Additionally, one Dutch child also followed this acquisition order. In a cross-linguistic study, Yavaş et al. (2008) showed that children acquiring Germanic languages (English, Norwegian, and Dutch) were more accurate at producing branching onsets, as opposed to children acquiring Hebrew who showed greater accuracy on sC clusters. Curiously, Carlisle (2006) reported that second-language learners acquire sC clusters before branching onsets (Carlisle 2006).

Contra Yavaş et al. (2008), Gierut's (1999) findings showed the opposite results in English: some children mastered sC clusters before branching onsets. Within the class of sC clusters, Fikkert (1994) reported that *s+stop* clusters were not acquired at the same time as *s+sonorant* clusters and branching onsets. She suggested that this confirms the assumption that only *s+obstruent* clusters have different properties. Hence, the research shows that there are cross-linguistic discrepancies in the order of acquiring sC clusters, and possibly even variation within specific languages.

Regarding the matter of simplification patterns, some studies, on the one hand, showed that there are no discrepancies between sC clusters and branching onsets. In the data from one child acquiring American English discussed by Gnanadesikan (2004), sC clusters were reduced to the less sonorous consonant, in the same fashion as branching onsets. The initial /s/ was deleted in the *s+stop* clusters (6a). In *s+sonorant* it was the second, more sonorous consonant, which was deleted (6b). Interestingly, this pattern mirrors onset reduction in reduplication occurring in Sanskrit (Gnanadesikan 2004).

(6) sC cluster reduction in English (from Gnanadesikan 2004, p.78)
  a. /skaɪ/ → [gay], ‘sky’
    /spuːn/ → [bun], ‘spoon’
  b. /snoʊ/ → [so], ‘snow’
    /sliːp/ → [sip], ‘sleep’

Other studies revealed that certain sC clusters are targeted by different simplification processes to branching onsets. In the study conducted by Krämer et al. (2017), most children acquiring Latvian reduced sC clusters to the least sonorous consonant, alike branching onsets (7a). However, *s+nasal* clusters did not follow this pattern and frequently the more sonorous nasal was retained in produced onsets (7b).

(7) sC cluster reduction in Latvian (adopted from Krämer et al. 2017, p. 9)
  a. /sle:dz/ → [se:dz] ‘he/she closes’
    /spainis/ → [painis] ‘bucket’
  b. /snieks/ → [nieks] ‘snow’

Smit (1990) reported similar results for English: /sl/ clusters were reduced to /s/, while in clusters with a stop, nasal or glide in the second position /s/ was deleted. Nonetheless, in the empirical inquiry of Yavaş et al. (2008) described above there was a slightly different tendency documented in all four languages that he investigated. Clusters consisting of *s + nasal* and *s + stop* were most frequently reduced to the second consonant, and /s/ was deleted. In /sl/ clusters there was a strong tendency to retain /s/, as also reported by Krämer et al. (2017) and Smit (1990). However, contrary to Smit’s findings, *s+glide* clusters patterned with *s+liquid* clusters in the retention of /s/. All the discussed findings suggested that *s + nasal* clusters, and possibly *s+glide* clusters most frequently undergo the reduction pattern to the less sonorous consonant and, thus, patterns together with branching onsets.

To summarize, Fikkert's (1994) study suggested that all *s+sonorant* sequences, contrary to *s+obstruent* clusters, have the same structural properties as branching onsets. Nevertheless, findings from studies investigating simplification patterns did not confirm this prediction since s+*nasal* patterned with *s+stop*. In both cluster types the initial /s/ was typically deleted, and *s+nasal* clusters did not follow the sonority pattern of reduction. It might indicate that in *s+stop* and *s+nasal* clusters, /s/ is not a part of a complex onset, but rather an extrasyllabic element (Goad 2011). Some studies also revealed significant differences within and across languages regarding the order of acquisition. In the following subsection, I discuss the effect of language-specific frequencies on the acquisition, which can possibly explain the observed discrepancies.

2.3 The role of input frequency in phonological acquisition

Even though segments are often distributed within syllables based on their sonority profiles, many languages have SSP-violating syllables in their inventories. These syllable types occur with varying frequencies across different languages. Recent research on phonological development showed that the frequency with which certain structures and segments occur in a language can influence the order of acquisition. It might be intuitive that the more frequently children hear a speech sound or a structure, the more chances they will have to associate it with its perceptual correlates. When a speech sound occurs more frequently in the ambient language, children will have more opportunities to master it (Everett & Schwartz 2023, p. 6).

In the first place, cross-linguistic comparisons revealed that there are differences regarding the age by which single phonemes are acquired across languages (Edwards et al. 2015). Several studies found that the order of acquiring consonants can be influenced by single phoneme and phoneme-sequence frequency in the ambient language (e.g., Beckman & Edwards 2010, Romani et al. 2017). The same results were found by studies that investigated the frequency effect on the acquisition of syllable types. Jarosz et al. (2017), using observations based on the spontaneous speech of four children acquiring Polish, reported that the acquisition of syllable types can be influenced by input frequency. In this study, the frequencies of different syllable types were calculated based on a corpus that contained the speech of caretakers interacting with the children. The results showed that frequency influenced accuracy in the production of syllable types. They also found that different frequency measurements led to different predictions (Jarosz et al. 2017, p. 384). The authors reported that *type frequency*, which relates to the occurrence of a given structure in the lexicon, made less accurate predictions than *token frequency* which relates to the total occurrence of these syllable types in the corpus. On the other hand, the findings of Beckman & Edwards (2010) supported the role of type frequency in two typologically distinct languages: English and Cantonese. The results of this study showed that accuracy in producing consonants is correlated with the frequency of their occurrence in the lexicon of the ambient language and the effect of type frequency overrode the effect of token frequency. For instance, the low-type frequency phoneme /ð/, which is acquired relatively late in English, has a significantly higher token frequency.

Zamuner et al. (2004) found that another input-frequency measurement of syllable types, known as *phonotactic probability,* can make correct predictions regarding the production of codas in English. Phonotactic probability captures the frequency with which single phonemes and phoneme sequences occur in a certain position in words in a target language (Jusczyk et al. 1994, Vitevitch & Luce 2004). To illustrate, a cluster /nt/ is not a possible syllable onset in English and its phonotactic probability of occurring in the word-initial position is equal to zero. However, the cluster /nt/ can occur in a coda position in English, for example, in the word *paint,* so the combination of these two segments has a higher probability of occurring in the word-final position. Another study that investigated the impact of phonotactic probability on acquisition was conducted by Edwards et al. (2005). This study tested the accuracy of the production of nonwords in children acquiring English. These nonwords varied in phonotactic similarity to the native words. The results showed that two phoneme sequences with high phonotactic probability were produced more accurately compared to the sequences with low probability. Jusczyk et al. (1994) found that infants at the age of 9 months could recognize nonwords that were composed of frequent sounds in their native language. This study also demonstrated that phonotactic probabilities calculated based on a corpus with adult-directed speech exhibited the same effect as probabilities based on child-directed speech. The authors argued that a similar frequency effect will be detected regardless of the source of input. On the other hand, it has been widely discussed that child-directed speech significantly differs from adult speech mainly in the hyper-articulation of target words (e.g., Kuhl et al. 1997, Ludusan et al. 2021). Therefore, it remains debatable whether using adult-speech corpora as frequency measurements in acquisition research gives the same outcomes as using child-directed corpora.

Regarding the acquisition of sC clusters, there are no studies that thoroughly investigate the relationship between the input frequency and the order of acquiring sC clusters. Yavaş (2014) observed that children acquiring languages with rich cluster inventories, which include both sonority plateaus and rises, exhibit different tendencies in the accuracy of sC cluster production than children acquiring Germanic languages. The authors of two studies investigating the acquisition of sC clusters in Norwegian (Kristoffersen & Simonsen 2006) and Dutch (Gerrits & Zumach 2006), only mentioned that frequency might impact the acquisition of initial clusters. However, the correlation between these two factors was not investigated closely. The findings of another study, conducted by Marecka & Dziubalska-Kołaczyk (2014), demonstrated that contrary to SSP predictions Polish children showed high productive accuracy for the *s + stop* cluster. Once again, the authors only suggested that frequency may play a role since /st/ is the third most common initial onset cluster in Polish.

Nonetheless, several studies showed that in some cases the frequency effect fails to predict the acquisition outcomes. One such study was conducted by Jarosz (2017) and investigated the acquisition of complex onsets in Polish, a language that has in its inventory many SSP-violating clusters comprising sonority reversals and plateaus. Type frequency of initial onset clusters for this study was calculated based on a frequency dictionary of child-directed speech. In Polish child-directed speech, plateaus composed of *obstruent* + *obstruent* sequence were more frequent than preferred sonority profiles such as *obstruent* + *liquid* and *obstruent* + *glide.* Nonetheless, children were more accurate in producing clusters with big sonority rises. Jarosz (2017) also demonstrated that computational models of phonological acquisition, which predict acquisition outcomes based on input statistics, did not capture low production accuracy of sonority plateaus. Likewise, frequency-based models presented in the study of Jarosz et. al. (2017), mentioned previously in this section, failed to capture low accuracy on SSP-violating onsets and articulatorily difficult segments, such as affricates, stridents, and trills. Both studies concluded that these results suggested that there might be an additional mechanism which drives phonological acquisition and named two potential factors: universal biases towards certain structures, and articulatory difficulty. This observation is in line with the Interaction Thesis proposed by Ambridge et al. (2015), originally for the acquisition of morphosyntax, and adopted by Edwards et al. (2015) for phonological acquisition. The Interaction Thesis states that sensitivity to input frequencies will interact with other factors during the course of acquisition. Edwards et al. (2015, p. 307) proposed that for phonological development *"…effects of frequency will interact with universal constraints on production and perception."* Thereby, the studies discussed demonstrated that language-specific frequencies sometimes fail to account for certain tendencies in phonological development. Thus, according to the predictions of the Interaction Thesis, both factors grounded in the target language and universal tendencies will shape the acquisition of the phonological system. Nevertheless, the question that arises is to what extent each factor influences the course of acquisition, and how these factors interact with each other. The study presented in this thesis will attempt to shed light on this question.

2.4 Theoretical assumptions about phonological acquisition

In this section, I discuss theoretical approaches to the first language acquisition of phonology. Firstly, I discuss the approaches adopting the theory of generative grammar, which captures the notion of markedness in phonological acquisition. The central assumption of generative grammar states that humans have an innate linguistic capacity - known as Universal Grammar - that plays a role in language acquisition. The first aim of generative acquisition theories is to establish properties of Universal Grammar, which enable a grammar to be acquired. Following this, the model needs to show how the grammar of each language can be learned by a child (Kager 1999). The discussed theories are linked to the process of acquiring syllable structure and consonant clusters, relevant to the current study. The generative approaches will be contrasted with cognitive- and perception-based approaches, which reject innate biases for SSP. Additionally, a reference is made to a recently introduced phonological theory, known as Substance-Free Phonology. Lastly, the perceptual-based approach Licensing by Cues is discussed in more detail as a possible account for the acquisition order of clusters.

2.4.1 Markedness and continuity in first language acquisition

The concept of markedness in phonological theory relates to the naturalness of phonological units and patterns. As discussed in the previous section, markedness can refer to cross-linguistic typological observations. However, it can also take a language-specific course and relate to the frequency of occurrence in a particular language (Rice 2007, Hume 2011). At the initial stage of acquiring grammar, children start to produce segments and structures that are commonly found in languages – therefore, being less marked - which was first observed by Jakobson (1948/1968). Based on this observation, Jakobson (1971) proposed that markedness shapes the order of language acquisition in the same way that it predicts implicational relations in language inventories. Thus, less marked segments and structures will be acquired before the corresponding marked categories. To illustrate, it has been broadly reported that the first syllables uttered by children are comprised of a single consonant and a vowel (CV), which is the most typologically preferred syllable type (Levelt et al. 2000). This is also related to the order of acquiring singleton consonants. As discussed earlier, in many languages, children acquire plosives before they acquire fricatives. Plosives are more frequent cross-linguistically than fricatives, and most languages with fricatives also have plosives, but not vice versa (Ladefoged & Maddieson 1996). A recent large-scale study, conducted by Everett & Schwartz (2023), confirmed this hypothesis. The study showed that a phoneme's typological frequency is more accurate in predicting the order of acquisition of consonants in English than language-specific frequency measurements. The results showed that bilabial sounds /p/ and /m/, which are common across languages, are acquired relatively early, even though these phonemes do not occur frequently in English.

It has been argued that markedness plays a role in both developing and mature grammar. *The Emergency of the Unmarked* is the notion that states that even though languages frequently have marked structures in inventories, the unmarked structures often "emerge", for instance in reduplication patterns (McCarthy & Prince 1994). Therefore, in generative theories, it is assumed that there is a continuity between child and adult language and all patterns found in the acquisition will reflect patterns occurring in languages (Pinker 1984). As mentioned earlier, Gnanadesikan (2004) drew a parallel between the onset reduction patterns found in child grammar and those in Sanskrit reduplication: both reflected the same onset reduction pattern. Sanskrit allows onset clusters, but, in reduplication, clusters are reduced to singleton onsets since the marked structures are not allowed in the reduplicant. Gnanadesikan's proposal and several other acquisition theories adhere to Jakobson's proposal and state that the same universal principles that shape phonological development are still active after the grammar is fully acquired (e.g., Fikkert 1994, Pater 1997, Goad & Rose 2004). Applying the same theory to syllable structure and onset clusters, if proposals regarding sonority-based markedness make correct assumptions about typological patterns, these generalizations will also predict which complex onsets and codas are the first to be acquired based on their sonority profile.

2.4.2 Language acquisition in Optimality Theory

Optimality Theory (Prince & Smolensky 1993/2004) is a theoretical framework, which accounts for adult and child grammar with the same principles. Optimality Theory (henceforth OT) is a constraints-based approach, and its central assumption is that grammar consists of a set of universal constraints. The actual pronunciation of words, known as a surface form, is labelled as output, and it is derived from the underlying representation, labelled as input. In the OT model of grammar, there are two types of universal constraints: markedness constraints, which are responsible for the well-formedness of the output, and faithfulness constraints, which account for the faithful mapping from inputs to outputs. Markedness constraints, which are claimed to be innate, must be motivated by typological observations (Krämer 2018). Constraints can be violated, and the output of the grammar results from solving a constraint conflict. Languages differ based on how constraints are ranked with respect to one another (Prince & Smolensky 1993/2004, McCarthy 2007). Therefore, certain constraints might not have a big impact on a particular language due to its relatively low ranking (Fikkert & de Hoop 2009). With regards to acquisition, the task of a learner is to discover the correct constraint ranking of the ambient language (Kager 1999). While acquiring grammar, children receive their auditory input from the adult’s surface form, from which they deduce the underlying form. In most acquisition models within the OT framework, the underlying representation of lexical items is assumed to be target-like. That is, it is assumed that children store segmentally accurate input forms (Boersma & Levelt 2003).

Since children produce words that are less marked than target words at the beginning stage of acquiring a language, most OT-based learning models assume that at the initial stage, markedness constraints outrank faithfulness constraints (e.g., Gnanadesikan 2004). A tableau presented in (8) shows a toy grammar that illustrates how high-ranked markedness constraints can account for the CV syllable shape at the early acquisition stage (based on the example from Fikkert & de Hoop 2009, p. 312).

(8) An OT tableau exemplifying high-ranked markedness constraints in child grammar.

Input***:*** /sneɪk /	NOCODA	*COMPLEX	ONSET	MAX-IO
a. sneɪk	*!			
b. sneɪ		*!		
c.☞neɪ				**
d. eɪ			*!	

NOCODA and *COMPLEX are two markedness constraints, which prohibit syllables with codas and consonant clusters, respectively. They are high-ranked, and their violation causes the elimination of candidate *a* and candidate *b*. Candidate *d* is eliminated by the violation of ONSET, which forces syllables to have an onset. The faithfulness constraint MAX-IO is violated by the most optimal candidate *c* since two segments from the input are missing in the output form.

After acquiring simple CV syllables, children gradually start to produce more marked syllable shapes with complex onsets and codas. Therefore, the constraint ranking shifts and NOCODA, *COMPLEX and ONSET are no longer ranked as highly. In this manner, most OT-based learning algorithms assume that learning is based on constraints re-ranking, and it is error-driven. When a learner notices a mismatch between the produced output and the target output, the grammar is re-evaluated, and the ranking of the constraints must shift for learning to progress (Tesar & Smolensky 1998).

It has been proposed that sonority-driven cluster simplification is caused by markedness constraints, which require well-formed onsets (Gnanadesikan 2004). To account for discrepancies in the reduction of sC clusters, Pater & Barlow (2003), based on data from children acquiring American English, argued that different patterns of reductions are caused by the conflict between constraints reinforcing low sonority onsets and constraints related to obstruents stopping. On the other hand, Goad & Rose (2004) emphasized that it is necessary to consider the structural distinction between the sC clusters and branching onsets. They argued that in addition to the constraints favoring low-sonority onsets, there is an active markedness constraint against an appendix at the edge of a syllable. This constraint accounts for a reduction pattern in which the initial /s/ is deleted in the output, regardless of the sonority profile of the following consonant. Further, Goad & Rose (2004) proposed three stages in acquiring consonant clusters. In the first stage, children do not produce clusters; the only possible outputs are singleton onsets. In the second stage, all clusters will be reduced to the least sonorous member. This stage is labeled as the *sonority pattern.* In the last stage, labeled as the *head pattern*, complex onsets are reduced to only the head of the clusters. The structural head of a cluster corresponds to the head of a syllable constituent. Branching onsets are assumed to be left-headed, so the first member of a cluster is the structural head (Kaye et al. 1990). Hence, in branching onsets, children delete the second consonant and retain the head of the cluster. On the other hand, sC cluster are assumed to be right-headed, since /s/ is supposedly an extrasyllabic element, or an appendix (e.g., Steriade 1982, Goldsmith 1990). Thus, extrasyllabic /s/ will be deleted from the sC cluster regardless of the cluster's sonority profile. Furthermore, they argue that developmental stages reflect attested typologies of languages with consonant clusters in their inventories. Certain languages, such as Spanish, do not allow sC clusters because of highly ranked constraints against the appendix but allow branching onsets, indicating that *COMPLEX is ranked low. Acoma allows sC clusters but not branching onsets. Therefore, Acoma's grammar has the opposite ranking, i.e., *COMPLEX ranked high, and

*APPENDIX ranked low (Barlow 2001, Goad & Rose 2004, p. 132). Thus, their analysis confirmed the continuity assumption.

2.4.3 Limitations of Optimality Theory

The “classical” version of OT has been criticized and questioned because this model of grammar could not capture variation occurring in the empirical data. However, there are many alternative constraint-based grammar algorithms that, for instance, let constraints float along a continuous scale (Stochastic OT, Boersma 1998), add noise to the evaluation (Gradual Learning Algorithm, Boersma & Hayes 2001), and allow variation caused by the non-grammatical factors such as frequency and speech rate (Noisy Harmonic Grammar, Coetzee 2016). In a model proposed by Anttila (1997), constraints are partially ordered. In the evaluation with Partially Ordered Constraints, the constraints are not ranked with respect to one another in each set, which assumes a random order for every evaluation. Therefore, the variation is not stipulated, but it is predicted by the model.

Even though OT-based acquisition models often assume that children’s underlying representation is segmentally accurate, it has been argued that this postulate is too simplistic (Fikkert & de Hoop 2009). Therefore, Boersma (1998) proposed that there are separate grammars for comprehension and production. In favor of the approach that does not distinguish between these two grammars, several studies show that children perceive the phonemic contrast of their native language before they start to speak, and that an infant’s babbling can reflect language-specific patterns (eg., Jusczyk et al. 1994, Boysson-Bardies et al. 1986). On the other hand, current research in cognitive psychology shows that the relationship between perception and production is not straightforward and might be independent. For instance, a study conducted by McAllister Byun & Tiede (2017), which investigated the acquisition of rhotic sounds in American English, confirmed that higher perception can anticipate the acquisition of this phoneme. Therefore, some speech errors can be associated not only with immaturity of the vocal tract, but also with lower perceptual skills. Smith (2010) highlighted that mispronunciation of words during the process of phonological development might be caused by both perception and production factors, and raised a question of whether acquisition theories should emphasize on competence (mental representation), or performance (actual pronunciation and perception). Further, he emphasized that the aim of the theory should be to explain patterns observed during the acquisition course. Therefore, a plausible theory of acquisition must consider both competence and performance. Smith (2010) proposed that children derive their own mental representation by passing adults' pronunciation through the perceptual filter. Overall, children's mispronunciation is the effect of overcoming limitations caused by physiological development (Smith 2010, p. 123-126). Thus, Smith's proposal regarding perception is aligned with recent findings regarding the relationship between perception and production. Nonetheless, Smith (2010) argues that errors occurring in child language are related to restricted performance, rather than competence.

The continuity assumption has also been challenged, mainly by consonant harmony. Consonant harmony occurs frequently in child language but does not occur in developed grammars, and it has been argued that only child-specific constraints can account for this process (e.g., Pater 1997). The approaches discussed in the previous subsection, which account for reduction patterns, adopt the strong position that all constraints are innate and universal without accounting for language-specific frequencies. However, there are also different proposals regarding the nature of constraints. The usage-based approaches to language learning (also labelled *cognitive approaches)* argue that constraints, among other aspects of grammar, are not innate for humans but are learned during the exposure to and usage of the language, i.e., through linguistic experience (Bybee 2010). From this standpoint, linguistic principles formalized as constraints are learned by general cognitive strategies which enable humans to make generalizations (van de Weijer 2017). This assumption is also adopted by Boersma (1998) and Hayes (1999), who reject the innate nature of constraints. To illustrate this point, van de Weijer (2017) demonstrated how *COMPLEX can be deduced by a child acquiring English. Only 3% of the most frequent words in child-directed speech contain consonant clusters. Thus, the preference for certain syllable shapes based on their sonority profiles would be deduced based on the data to which a learner is exposed. Additionally, Daland et al. (2011) demonstrated that a computational model can capture the preference of English speakers for SSP in nonce words based on lexical statistics.

Returning to the concept of the OT framework: this theory makes a strong reference to phonetic properties, known as *substance.* This assumption has also been questioned. A recent theory of Substance-Free Phonology argues that grammar is not shaped by substance, and phonological computation does not reference acoustic and articulatory correlates (Chabot 2022 for an overview). Thus, the substance-free approach does not refer to markedness, and the link between linguistic typology and child phonology is irrelevant in this light (Reiss 2018). Reiss (2018, p. 431-432) also highlights that the Emergence of the Unmarked in developing grammar might occur due to factors affecting performance rather than competence and argues that "*[t]he* *mental representation of a mouse is not smaller than that of an elephant. (…) and even if we demonstrate that producing [p'] is more physically challenging than producing [p], it does not follow that the mental representation of the former is more complex than that of the latter and needs to be fixed".* Thus, Reiss stated that that the complexity of marked segments is not related to the mental representation of grammar but rather to the physical complexity of pronouncing certain units. Further, he claims that posting innate constraints against certain structures might be misleading for the speakers whose language has these structures in the inventory, and unnecessary for the speakers whose language lacks them. From this point of view, errors made by children are related to performance, rather than a mental representation of grammar, and therefore, investigating child language is not relevant to phonological theory.

However, substance-free approaches do not completely abstain from innateness. For instance, Reiss & Volenec (2022) argued that humans possess an innate concept of phonological features and the capacity to map phonetic content onto features. These features, which partake in phonological computation – which itself is not dependent on the featural content – are innate for humans. Scheer (2022) argues that sonority is innate and phonologically meaningful, contrary to other features with phonetic content such as place and voicing. Namely, following Scheer (2022), sonority is not interchangeable and has cross-linguistically stable properties. Thus, sonority primes have innately associated phonetic correlates and the effect of sonority, contrary to other phonological features, will influence the phonological system. Nonetheless, Scheer's (2022) proposal does not address consonants clusters, and it mostly relates to the distinction between consonants and vowels.

2.4.4 Licensing by Cues

An alternative proposal to phonotactics, that does not adhere to sonority, or any sonority-based generalization, was initially proposed by Steriade (1999, 2009), and is known as Licensing by Cues. This account is based on perception, and it refers to strings of segments, rather than syllables, in explaining phonotactic behavior. Steriade (1999, 2009) argues that the perceptibility of contrasts in segments is based on their perceptual cues to the contrastive features. Contrasts can be licensed in the positions when acoustic cues can be recovered. When the acoustic cues are unrecoverable, the contrast is neutralized. That is, when two segments with contrastive features appear in one string, they are more likely to be perceived and licensed than segments with two similar features. Therefore, segments with contrastive features are more likely to occur adjacently within one string. These generalizations can be formalized by constraints in the OT framework (e.g., Steriade 2001). Henke et al. argued (2012) that Licensing by Cues makes better predictions than SSP regarding phonotactics and several other cross-linguistic tendencies, and that there is no need to propose innate bias towards sonority and against extrasyllabicity.

Vanderweide (2005) demonstrated how a perceptual account can capture the order of acquisition of segments and consonant clusters[1], and argued that markedness in linked to perceptibility, which shapes possible acquisition paths. This proposal is based on typological observations reported by Morelli (1999) who demonstrated that obstruent clusters with segments contrasting in the continuant feature e.g., *fricative* + *stop* are more frequently attested than clusters not contrasting in this feature e.g., *stop* + *stop.* According to Vanderweide's proposal, which also builds on other work regarding acoustic perceptibility, segments have internal and contextual acoustic cues. *Internal cues* refer to the acoustic robustness of segments in terms of the presence and type of continuous airflow and the degree of openness of the vocal tract. *Contextual cues* refer to cues generated when articulatory overlap occurs between two segments, and the strength of these cues depends on the differences in internal cues between adjected segments. Ordering segments based on their degree of constriction resembles the order stated by the Sonority Hierarchy (Henke et al. 2012, p. 74).

Vanderweide (2005) states that children will first acquire segments that are most acoustically robust in the most contrastive environments and thus, have the greatest perceptibility. By the way of illustration, this approach can explain the preference for the CV syllable shape over the VC shape, and the tendency towards reducing onset clusters to stops. Vowels are the most acoustically robust segments because they are articulated with the greatest openness and airflow, contrary to stops. However, stops are uttered with a release burst, and context of the greatest perceptibility for stops to occur is when a plosive is followed by a vowel. If a plosive is followed by a segment articulated with a narrower oral closure, the release burst is more likely to be lost and make the stop harder to perceive (Henke et al. 2012). Thus, a *vowel* + *stop* sequence has the most robust internal and contextual cues.

---

[1] Vanderweide (2005) refers to consonant clusters as *sequences* since the perception-based proposal does not refer to syllables.

Vanderweide (2005) proposed that the acquisition of CCV sequences will be constructed by the strength of aperture cues and the robustness of periodicity cues to the following consonant. The possible acquisition paths depend on whether aperture or periodicity cues are extracted by a child from the input, and therefore, variation in the order of acquisition will occur. To illustrate the concept of aperture cues, stops are audibly released before approximants which are articulated with continuous airflow. However, when a plosive occurs before a nasal, it is articulated silently since both segments are articulated with complete oral closure. For a *plosive + nasal* sequence, the robustness of periodicity cues to a following segment influences whether contextual cues can be recovered from the articulatory overlap. Nasals have weaker internal cues than approximants, and when they occur before an obstruent their periodicity cues are weaker. Table (3) presents two possible acquisition paths, depending on whether periodicity or aperture cues are extracted, proposed by Vanderweide (2005, p. 148-155)

	Stage I	Stage II	Stage II
Path 1 ***Aperture Cues***	***stops + liquid*** *fricative + stop*	*Stop + glide* *Fricative + stop*	*Stop + nasal* *Fricative + nasal*
Path 2 ***Periodicity cues***	***stops + liquid*** *fricative + approximant*	*Stop + glide* *Fricative + fricative*	*Stop + nasal* *Fricative + glide*

Table 3. Developmental stages in acquiring CCV (Vanderweide 2005).

Vanderweide (2005) argues that a stop followed by a liquid will be acquired first since liquids exhibit the greatest contrast in both openness of the vocal tract and the continuity of the airflow. Therefore, in this position, both segments have the strongest periodicity cues and will be the easiest to perceive. Stops, when followed by approximants, will be licensed before obstruents following nasals where the acoustic overlap is less robust. Liquids in the second position are favored over glides since glides have a more similar degree of openness to the following vowel. Further, there will be more variation in the acquisition of fricative-initial sequences than stops-initial sequences because fricatives in this position differ more than stops in the degree of the robustness of both cues.

Nevertheless, Vanderweide (2005) does not address the order of acquiring prevocalic sequences with sibilant fricatives and whether their acquisition diverges from the acquisition of other fricatives. Henke et al. (2012), also referring to Morelli's survey (1999), pointed out that *s +*

stop clusters are the most attested among all *fricative + stop* sequences. They claimed that this is due to more robust internal cues of sibilant, which makes sibilants to be more easily perceptible at the word-initial position before a plosive. This also explains why *s + stop* clusters are common in languages, even though this group of clusters violates the SSP generalization. Henke et al. (2012) argue that this explanation is more plausible than treating *s* as an appendix and stating extrasyllabicity since it does not require the exceptional status of these phonemes. In the light of Vanderweide's (2005) perception-based account for the acquisition order, segments with the more robust cues will be acquired first. Thus, sC clusters containing stops as the second member will be easier to perceive for children than clusters with other fricatives as followed by a stop, and therefore, are first to be acquired.

2.5 The scope of the current study

The main objective of the study conducted in this thesis is to contribute to research on phonological acquisition and investigate factors underlying the process of acquiring consonant clusters. I investigate whether two sonority-based generalizations: the Minimal Sonority Distance, and the Sonority Dispersion Principle, can make correct predictions about the order of acquisition of biconsonantal onset clusters in Icelandic. Additionally, I investigate the simplification patterns that target initial onset clusters. As discussed in subsections 2.2.2. and 2.2.3, there is a discrepancy between the phonological processes that target branching onsets and sC clusters. Branching onsets are usually reduced to the least sonorous member of the cluster, while sC clusters are not targeted by sonority-driven simplification. However, other studies demonstrated that only *s* + *stop* and *s* + *nasal* clusters do not follow this pattern, while *s* + *approximant* clusters often undergo sonority-driven reduction (Yavaş et al. 2008, Krämer et al. 2017). The current study aims to investigate whether this observation will be confirmed by data from children acquiring Icelandic. The differences in simplification patterns found in child data can contribute to an understanding of whether sC clusters are not structurally tautosyllabic, but rather that in certain sC clusters the initial /s/ should be analyzed as an extrasyllabic element (Goad & Rose 2004, Goad 2011). If assumption of the special status of sC clusters does not hold true, there will be no difference in phonological processes targeting these group and branching onsets.

The second aim of this study is to investigate whether the data from Icelandic children confirm the predictions of the frequency effect and sonority-based generalizations about the process of phonological acquisition. As discussed in subsection 2.3., the frequencies with which certain structures occur in the ambient language can impact the order of acquisition of segments and consonant clusters. However, the studies conducted by Jarosz (2017) and Jarosz et al. (2017) have shown that frequency in isolation does not always make correct predictions regarding the order of acquisition in Polish. Moreover, Jarosz (2017) revealed that children produced clusters with bigger sonority distances more accurately, even though the majority of their input contained sonority plateaus. The authors of both studies suggested that phonetic difficulty or universal biases towards certain structures may influence the acquisition order. In light of their findings and the conclusions, I investigate the effect of frequency, sonority-based markedness, and articulatory difficulty on the order of acquisition of complex onsets in Icelandic. According to the predictions of the Interaction Thesis (Ambridge et al. 2015, Edwards et al. 2015), both frequency and universal tendencies will impact phonological acquisition. Furthermore, in the current study, I attempt not only to investigate which factors underlie the acquisition of consonant clusters, but also - by the application of statistical analysis - to show which of the investigated factors can make the best predictions about the order of acquisition.

According to the assumption of acquisition made in OT, children produce unmarked structures at the initial stage of acquisition. According to OT predictions about phonological acquisition, the least marked clusters will be acquired first, and gradually, over the course of acquisition children start to produce more marked structures. Regarding initial tautosyllabic clusters, there are two principles which classify them by their complexity: the Minimal Sonority Distance (henceforth MSD) and the Sonority Dispersion Principle (henceforth SDP), discussed in subsection 2.1.3. Conversely, Kistanova (2021) showed that sonority cannot predict the order of acquiring consonant clusters. Kistanova (2021) also suggested that articulatory difficulty and frequency may affect the course of acquisition. In the study conducted in this thesis, I test the predictions of these principles of the MSD and the SDP to investigate whether they validated by children's data and can predict the order of acquisition. In the current study, the effect of sonority will be contrasted with the effect of the articulatory difficulty of segments. As reported by studies investigating the order of acquisition of singleton consonants in many languages, children acquire stops before fricatives (McLeod et al. 2001). Fricatives are judged to be more difficult to articulate. Hence, if the articulatory difficulty of the individual consonants in the cluster is predictive of the order of acquisition, stop-initial clusters should be acquired before fricative-initial clusters. It has been noted that, in Icelandic children acquire the lateral /l/ before the trill /r/. If there is an effect of the articulatory difficulty of individual segments, the clusters with /r/ will be acquired after the clusters with /l/ as a second member.

Lastly, the findings will be discussed with respect to the perception-based approach. As described in subsection 2.4.4., the perception-based proposal Licensing by Cue disregards the role of sonority in shaping the organization of segments within the syllables, and attributes it to the role of perceptual cues. According to Vanderweide's (2005) proposal, the contrast in *stop + liquid* clusters have the strongest perceptual cues and therefore, *stop + liquid* is the first cluster type to be acquired. As proposed by Henke et al. (2012), *s + stop* clusters are common in languages, even though this class violates SSP, because /s/ is more likely to be perceived before a stop, than before a fricative. To investigate all the described factors, I analyze the dataset which includes words with initial onset clusters produced by children acquiring Icelandic. Before proceeding to the chapter that elaborates on the research questions, describes methodology and presents the analyzed data, I discuss the classification of Icelandic consonants into sonority classes which is essential for investigating the role of sonority in acquisition.

2.6 Phonology of Icelandic

In this section, I discuss the phonological properties of Modern Icelandic with a focus on consonant inventory and syllable phonotactics. Lastly, I summarize findings regarding the acquisition of initial onset clusters in Icelandic.

2.6.1 Consonantal inventory

Icelandic belongs to the family of the North Germanic languages together with Norwegian, Danish, Swedish and Faroese. Unlike other Nordic languages, Icelandic has rich inflectional morphology and a distinctive phonological system (Árnason 2011). As noted by Másdóttir et al. (2021), Icelandic dialects have minor differences regarding phonology. A consonantal inventory of Icelandic is presented in table (4). Icelandic has 18 phonemic consonants in its phonological system. There are 10 allophonic consonants that do not contrast phonemically, which are given in table (4) in parentheses.[2]

Manner	**Place of articulation**					
	Labial	Dental	Alveolar	Palatal	Velar	Glottal
Stops	pʰ p		tʰ t	(cʰ) (c)	kʰ k	(ʔ)
Nasals	(m̥) m		n̥ n	(ɲ̊) (ɲ)	(ŋ̊) (ŋ)	
Fricatives	f	θ	s	ç	(x)	h
Approximants/ Voiced fricatives	ʋ/v	(ð/ð̞)		j/ ʝ	ɣ/ ɰ	
Laterals				l̥ l		
Trill				r̥ r		

Table 4. Icelandic consonants (based on Árnason, 2011, p. 98, Másdóttir & Stokes 2016, p. 2)

---

[2] It must be noted that sometimes all voiceless sonorants are treated as allophones of their voiced counterparts (e.g., Heimisdóttir 2015). Since these phonemes do not occur in complex onsets, this issue will not be discussed further.

All Icelandic stops are voiceless, and they contrast in aspiration (Árnason 2011). Voiced approximants traditionally have been described as fricatives ([v], [ð], [ʝ], [ɣ] respectively), and some scholars treat these phonemes as such. However, Árnason (2011, p. 108) notes that they are frequently articulated as approximants or glides. Further, Árnason (2011, p. 164) demonstrated that /j/ and /v/ have similar phonotactic behavior to other sonorants. Phonemes /j/ and /v/ occur frequently as a second member of a complex onset, unlike any other fricative. Moreover, Másdóttir & Stokes (2016), referring to Helgason's (1991) proposal, treats these phonemes as approximants in their study investigating the acquisition of Icelandic consonants.

Heimisdóttir (2015), in an analysis of Icelandic aspiration, demonstrated in her dissertation that this class of phonemes has different phonotactic behavior to other fricatives. When aspirated stops and /s/ are followed by approximants /v, j, r/ word medially, these two phonemes syllabify as a complex onset, which is an exception to the preference for heterosyllabic clusters in Icelandic. Only clusters with the biggest sonority distance can be syllabified as complex onsets and thus, /v/ and /j/ have a greater sonority value than fricatives in Icelandic. Thus, Heimisdóttir (2015) treats both /v/ and /j/ as approximants. Considering the syllabification rules, Heimisdóttir (2015) proposed a sonority hierarchy of segments specific to Icelandic (Figure 5). Heimisdóttir (2015) also discussed aspirational alternations regarding internal clusters. An observation of note is that /v/, unlike any other phonemes from this class, can also appear as a first member of a complex onset and combine with /j/ in /vj/. Árnason (2011) suggested that this phonotactic restriction may imply that /j/ behaves more as an approximant, than /v/.

Most sonorous (6) *v, j*
(5) *r*
(4) *m, n, l*
(3) *ɣ, ð*
(2) *f, θ, x*
(1) *p, t, k*
Least sonorous (0) p^h, t^h, k^h, *s*

Figure 5. The Sonority Hierarchy specific to Icelandic (Heimisdóttir 2015, p. 21).

This also indicates that /v/ is less sonorous than /j/ and has properties which differentiate it from other sonorants in Icelandic (Bjorndahl 2018). The special status of /v/ has been discussed by many scholars, and it has been proposed that in some languages, /v/ can be specified for different values of the sonorant feature, which is determined by phonotactics (Chew 2003). For instance, in Russian /v/ should be analyzed as an obstruent in the initial position since it undergoes voicing assimilation to the following obstruent, and as a sonorant in the internal prevocalic position (Bjorndahl 2018). Thus, in Icelandic onset clusters /v/ might behave more as an obstruent, while when it occurs as initial consonant in a cluster as in /vj/, and when it occurs as a second member of a cluster, it might behave as a sonorant, as in /sv/.

In the current study, firstly, I present the predictions of the MSD and the SDP based on the Sonority Hierarchy specific to Icelandic and adapt the phoneme classification proposed by Heimisdóttir (2015). Secondly, I present the predictions of the MSD and the SDP based on the Sonority Hierarchy presented in figure (3). Here, based on patterns found in the data, /v/ will be treated as voiced fricative, while /j/ as an approximant in initial clusters. This classification will be further justified in section 5.4., where the data analysis showed that /v/ seems to pattern differently in initial clusters than other sonorants with respect to the MSD. Lastly, I show how children performed on different natural classes to assess whether /j/ and /v/ patterns with other approximants which occur as a second member of a cluster.

2.6.2 Initial consonant clusters

According to Thráinsson & Gíslason (1993, as cited in Másdóttir et al. 2021), Icelandic has approximately 70 consonant clusters. As described by Árnason (2011, p. 163-165), nasals and liquids /m/, /n/, /l/, /r/ can only combine with /j/ to form a complex onset. Voiceless fricatives /f/ and /θ/ can combine with sonorants e.g, /fj/, /fl/, /θv/, /θr/. There are also fricative-initial clusters composed of three consonants where *fricative + sonorant* is followed by /j/ e.g., /flj/. Initial stops, both aspirated and unaspirated, can combine with sonorants e.g., /pl/, /pʰj/, /tr/, /tʰv/, /kr/, /kʰl/. Stop-initial clusters can also combine with /j/ to form triconsonantal clusters i.e., /pʰrj/, /trj/. Regarding voiced approximants, only /v/ can appear word-initially, and it can only combine with /j/ in the cluster /vj/.

Icelandic has sC clusters in its inventory, and /s/ can combine with both sonorants and unaspirated stops /k/, /p/, and /t/ i.e., /sj/, /sv/, /sp/, /sk/, /st/. However, /s/ cannot combine with /r/ to form a complex onset. When /s/ is followed by nasals /m/, /n/ and lateral /l/, the cluster can be pronounced with an intrusive stop as /spm/, /stn/ and /stl/ respectively. Icelandic has triconsonantal clusters with the initial /s/ followed by a stop and a sonorant: /spr/, /str/, /skr/, /spl/, /skv/. Both bi- and triconsonantal sC clusters can combine with /j/, e.g., /spj/, /skrj/, /strj/.

In Icelandic, main stress falls on the initial syllable (Árnason 2011). Thus, initial onset clusters, investigated in the current study, always occur in the stressed syllable. This should be noted, because studies investigating factors influencing acquisition of onset clusters reported that a cluster that occurs in a stressed syllable may be produced more accurately than clusters in unstressed syllables (e.g., Jarosz 2017).

2.6.3 Previous findings on the acquisition of complex onsets in Icelandic

A longitudinal study conducted by Másdóttir (2008) investigated the phonological development of 28 Icelandic children between the ages of 2;4 and 3;4 (years; months). The study mainly investigated the acquisition of single phonemes, however, some findings regarding initial clusters were discussed. The results showed that initial sC clusters are produced with lower accuracy than stop-initial clusters. Másdóttir (2008) also noted that clusters /sm/ and /sn/ were hardly ever pronounced with an intrusive stop by children, contrary to /sl/ which was commonly produced as /stl/.

A large-scale study, conducted by Másdóttir et al. (2021), investigated the production of consonants and consonant clusters in 437 normally developing children acquiring Icelandic between the age of 2;6 and 7;11. The children were tested on singleton consonants and clusters with the Icelandic Speech sound test Málhljóðapróf ÞM (Másdóttir 2014, ÞM's Test of Speech Sound Disorders). To calculate the accuracy of production, a percentage of consonants correct (PPC) was calculated for both singleton consonants and clusters. The PCC score is commonly used by speech pathologists to assess the percentage of correctly produced consonants. It is calculated by dividing the number of correctly produced consonants/clusters by the total number of consonants/clusters (Shriberg et al. 1997). Using the criteria of 90% PCC for a cluster to be acquired, the results showed that /pl/ was the first cluster to be acquired at the age of 2;6-2;11. All *stop + sonorant* clusters were acquired at the age of 5;6-5;11. The sC clusters, /sn/ and /sj/ were acquired at the age of 4;6-4;11, and /sp/ and /st/ were acquired at the age of 5;5-5;11. The last clusters to be acquired, at the age 6;0-7;11, were /sm/, /sv/ and triconsonantal clusters /stl/, /str/ and /skr/. Also, the cluster /θr/ was acquired at this age. Overall, clusters were produced significantly less accurately than singletons. The findings also demonstrated that children acquire singleton consonants before clusters.

Regarding singleton consonants, Másdóttir et al. (2021) reported that /m/, /n/, /p/, /t/, /j/, /h/ are the first consonants to be acquired. A trill, voiceless sonorants, and fricatives were the last to be acquired. These phonemes were acquired around the age of 6;0-6;11. The phoneme /l/ was acquired relatively early by Icelandic children, at the age of 3;0-3;5. A longitudinal study, conducted by Másdóttir (2008), investigated the order of acquisition of consonants by 28 children acquiring Icelandic (aged 2;4-3,4). Categorizing by the manner of articulation, Másdóttir (2008) reported that at the first stage, Icelandic children acquired nasal, laterals, and stops, followed by fricatives and approximants. Children acquired trills last. This order of acquisition has been previously reported for Germanic languages. However, the phoneme /l/ is acquired relatively earlier than e.g., in English (McLeod & Crowe 2018).

3. Methodology

3.1 Research questions

The study conducted in this thesis aims to investigate the acquisition of complex onsets by children acquiring Icelandic from a theoretical perspective. The first aim of the study is to investigate reduction patterns occurring in the developing phonological systems of children acquiring Icelandic. The discussed studies, which investigated reduction patterns applied to consonant clusters, showed that branching onsets are usually reduced to the least sonorous consonant. However, studies investigating patterns applied to sC clusters reported that certain sC clusters seem to diverge from the reduction pattern driven by sonority. This divergence might reveal the structural differences between the two types of clusters.

The second aim of the study is to investigate factors influencing the accuracy of producing complex onsets. There are two major factors discussed in the literature that affect the developmental path: universal biases towards unmarked structures, and frequency with which certain structures occur in the ambient language. According to previous research on phonological acquisition, both factors influence the process of acquisition to a certain extent. Additionally, research conducted by speech pathologists suggested that children who mastered more singleton consonants will be more accurate when producing consonant clusters. The study conducted in the thesis will attempt to tease apart all the factors to examine which plays the biggest role. Thus, as stated in the Introduction, I address the following research questions:

1. Is there a difference between reduction patterns occurring in branching onsets and sC clusters in Icelandic?
2. Do the Sonority Dispersion Principle and the Minimal Sonority Distance make the correct predictions regarding the order of acquisition of initial onset clusters in Icelandic?
3. Does the type frequencies of initial consonant clusters in Icelandic affect children’s accuracy in producing initial onset clusters?
4. Did children who acquired more singleton consonants also acquire more initial onset clusters?
5. Is there a correlation between the order of acquisition of natural classes and the order of acquisition of initial onset clusters within these classes?

3.2 The Másdóttir Corpus

To investigate the addressed research questions, I analyze the data which are available as a part of the Icelandic Másdóttir corpus, which is a part of the PhonBank corpora (Rose & MacWhinney 2014). The Másdóttir corpus includes the speech production of 288 typically developing monolingual children acquiring Icelandic at the age of 2;6-7;11 (Másdóttir 2014, Másdóttir et al. 2021)[3]. The data were collected using tasks from the Icelandic Speech sound test Málhljóðapróf ÞM, elaborated upon by Másdóttir (2014). The test consists of three tasks in which children were tested for their level of accuracy in producing multisyllabic words, singleton consonants and consonant clusters, and consistency in producing the same words multiple times. During the test, the children were asked to name pictures, and the target words were elicited spontaneously. When a child was not able to produce the target word spontaneously, imitation was used. The data were transcribed by three experienced transcribers, and the Másdóttir corpus includes the transcription of words produced by tested children and transcriptions of the target forms. Additionally, the corpus includes audio recordings of each session. The data are grouped by ages of children, in 6-month age intervals for the youngest children (2;6-5;11), and in 12-month intervals for the oldest children (6;0-7;11).

3.3 Processing the data.

To test children's accuracy on clusters, I analyzed the data from the Malhjodaprof part of the test, which examined the production of singleton consonants and consonant clusters. In the Malhjodaprof part, children were assessed on 96 words that included 47 single consonants and 45 consonant clusters (initial, medial, and final). In this part of the corpus, in addition to the transcription of produced words and the target transcription, there is an annotation describing how many consonants were produced correctly in each word. The Malhjodaprof included 36 words with biconsonantal initial consonant clusters, and these words were included in the dataset for the main analysis. All words with their target transcription are included in Appendix I. In the analysis, I included only the subjects who systematically reduced consonant clusters. Thus, most of the dataset comes from the youngest children. This produced a dataset of 2411 words with initial onset clusters in the target form produced by 68 children (1413 observations from males, 998 observations from females) between the ages of 2;6 and 4;3 (Figure 6, M = 36.95 (months), SD = 5.33).

---

[3] Children with cleft/lip palate, or developmental disorders were not included (Másdóttir et al. 2021, p. 1493).

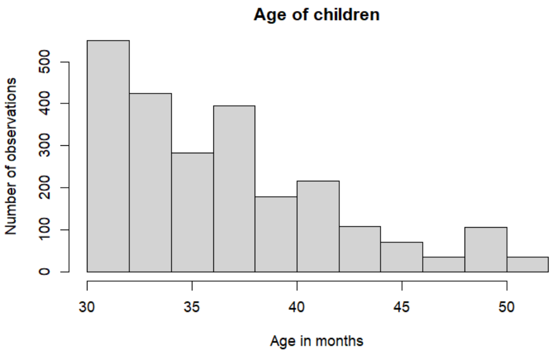


Figure 6. Age of children given in months.

Most children produced all 36 words, but some subjects failed to produce all of them. The number of observations from each child is included in Appendix I. There are 1609 observations of words with initial branching onsets and 802 observations of sC clusters. Initial biconsonantal onset clusters included in the dataset are summarized in table (5). Relative frequencies and raw frequencies of each cluster in the dataset are given in Appendix 1.

Cluster type	Onset clusters	Percentage of the total dataset
Branching onsets	**/fl/, /fr/, /θr/, /θv/,** **/lj/, /mj/, /pj/, /tʰv/** **/pl/, /kl/, /kr/, /pr/, /tr/** **/pʰl/, /kʰl/** **/pʰr/, /tʰr/, /kʰr/**	67%
sC clusters	**/sp/, /sc/, /sk/, /st/,** **/sn/, /sm/, /sv/, /sj/**	33%

Table 5. Initial onset clusters included in the dataset

As noted earlier in section 2.4.2, /sn/, /sm/, and /sl/ tend to be pronounced with an intrusive stop by Icelandic speakers. However, in the corpus, the target transcription for the *stop* + *nasal* clusters were annotated without intrusive stop. Also, no child produced these clusters with an intrusive stop, which is in accordance with the observation reported by Másdóttir (2008). There were only a few cases where /sn/ was produced as /tn/. In these cases, the target transcription was assumed to be without an intrusive stop in accordance with the Másdóttir corpus.

Additionally, the Malhjodaprof included four words with three different initial triconsonantal sC clusters (sCC): /stl/, /str/, /skr/. In total, 267 words with sCC clusters were analyzed based on the reduction patterns applied to them. However, this part of the data was not included in the main analysis, given that neither Sonority Distance nor Sonority Dispersion allows for measuring the complexity of triconsonantal clusters in relation to biconsonantal clusters[4]. Accuracy on triconsonantal onsets was also annotated to compare this group to the branching onsets and biconsonantal sC clusters. Words with sCC clusters, and their raw and relative frequencies are summarized in Appendix I.

The data were extracted into Excel and manually sorted by the subject number, age of children, sex, attempted initial clusters, produced initial clusters, simplification patterns applied to the clusters, and accuracy (0 indicating *inaccurate*, 1 indicating *accurate*). Additionally, the independent variables, which are addressed in the research questions, were included. The levels of all factors will be discussed in detail in the next subsections.

3.3.1 Accuracy and reduction patterns

The data analyzed in the current study data, have already been investigated in the study conducted by Másdóttir et al. (2021). However, as described in section 4.3., in this study the researchers used PCC as the measure of accuracy for both singleton onsets and consonant clusters. Thus, when both consonants in the cluster were produced, but segmental changes occurred, the cluster was counted as produced inaccurately[5]. In the current study, when two segments in a cluster were produced but segmental changes occurred due to other phonological processes occurring in child phonology, such as stopping, fronting, and gliding, the cluster was counted as produced accurately if the segments were substituted with consonants of the same manner of articulation. The same method of determining accuracy was applied in the study conducted by Jarosz (2017). All cluster simplification processes are demonstrated in Table (6).

---

[4] Even though, as demonstrated in section 1.3., Sonority Dispersion allows one to account for triconsonantal clusters, in Icelandic these are composed of two initial obstruents, followed by a stop or liquid. Nevertheless, Sonority Dispersion does not account for sonority plateaus and reversals, and all sCC clusters consist of a fricative followed by a stop.

[5] As noted in Másdóttir et al. (2021, p. 1495) only errors with /r/ were excluded in cases where /r/ was produced inaccurately in a grammatical suffix.

Phonological processes	Examples	Accuracy
1. No reduction	a. /klʏccɪ/ → [klʏccɪ], *a window* b. /tʰrɔmma/ → [tʰðɔmma], *a drum* /kreiːða/ → [kleiːða], *a comb* c. /kʰlʏhka/ → [tlʏhka], *a clock* d. /plauːsa/ → [tlouːta], *to blow* e. /scaiːrɪ/ → [θcaiːrɪ], *scissors* /sjouːɹ̥/ → [s̠jouːɹ̥], *a sea*	***accurate***
2. Reduction to the least sonorous consonant	a. /pʰrɪnsɛssa/ → [pɪːsɛssa], *a princess* /plouːm/ → [pouːm], *a flower* /frɔskʏɹ̥/ → [fɔhkʏ], *a frog* /kriːs/ → [tiːθ], *a pig* b. /stoutl̥/ → [tɔtl̥], *a chair* /spɪːl̥/ → [pɪːl̥], *a game* c. /sjouːɹ̥/ → [souːɹ̥], *a sea* /sviːn/ → [siːn], *pork*	***inaccurate***
3. Reduction to the most sonorous consonant	a. /ljouːn/ → [jounɪ], *a lion* b. /smɛhkʏɹ̥/ → [mɛhkʏɹ̥], *a bib* /snʏːθ/ → [nʏːθ], *a pacifier*	
4. Coalescence	a. /sviːn/ → [fiːn], *pork* b./θvɔhtavjɛːl̥/ → [fɔhtavɛːl̥], *a washing machine* c. /pʰlaustʏɹ̥/ → [l̥auhtʏ], *a band-aid*	
5. Metathesis	/stɛl̥pa/ → [tlɛppa], *a girl*	
6. Epenthesis	a. /pruː/ → [pəru], *a bridge* b./ mjoul̥ka/ → [mɔjɔhka], to *milk/milking*	
7. Deletion	/θriːɹ̥/ → [iːθ], *three*	
8. Other processes	a. /tʰveiːɹ̥/ → [fðeiːθ], *two* b. /snʏːθ/ → [tnʏːθ], *a pacifier*	

Table 6. Phonological processes that occurred in cluster simplification with examples.

The most common segmental substitution occurred with the trill /r/, which is one of the last consonants to be acquired in Icelandic, as Másdóttir et al. (2021) reported. In (1b), /r/ was usually substituted with other liquids or approximants. The case of /r/ substitution is the only example of a segment frequently being substituted with a consonant of a different manner of articulation. In this case, the clusters were counted as accurate since both segments were produced, instead of being reduced to one consonant, especially since /r/ substitution is a commonly reported pattern in children phonology. Similarly, both segments were produced in (1c), but the velar stop /kʰ/, produced in the back of the mouth, lost aspiration and was substituted with the alveolar [t] produced in the front of the mouth. This process is known as backing, and it frequently occurs in child phonology. In (1d), consonant harmony occurred, and initial /p/ was produced as [t] because of assimilation to /t/ that occurred later in the word. Another commonly occurring substitution was the replacement of /s/ by the dental fricative /θ/. Alternatively, /s/ was pronounced as dental [s̪] (1e). The productions that included these substitutions were still considered to be accurate.

A cluster was marked as inaccurate if only one segment was produced, or one or both segments were substituted with disparate segments, as presented in (8b). As mentioned earlier, there were a few cases where in /sn/ - which can be produced with an intrusive stop by Icelandic speakers - /s/ was substituted with a stop (8b). When a cluster was reduced to the least sonorous consonant, the obstruent was retained (sometimes with segmental changes), while the sonorant was deleted (2a). In sC clusters, that consisted of *s + stop* sequences, /s/ was deleted and a less sonorous stop was retained (2b). However, when this pattern occurred in /sv/ or /sj/, the less sonorous /s/ was retained (2c)[6]. In other cases, clusters were also reduced to the most sonorous consonants, as presented in (3a). When this reduction pattern occurred in *s + nasal* sequences, the nasal was preserved and the less sonorous /s/ was deleted (3b). Coalescence occurred when two segments were merged, and replaced with another segment that shares features of both consonants from the target cluster (4). For instance, in (4a) the produced segment [f] has the labial feature of /v/ and the voiceless fricative manner of /s/, while in (4c) the preserved /l/ gained the voiceless feature from /p/ and was realized as voiceless [l̥]. Metathesis occurred when two segments changed their order (5), epenthesis occurred when a schwa or other vowel was inserted between two consonants, and deletion where both segments were missing (7).

3.3.2 The measurements of sonority-based markedness

One aim of the study was to investigate the effect of markedness – as based on the sonority profiles of initial onsets – on accuracy in cluster production. However, studies on cluster

---

[6] Whether /v/ is treated as approximant, or as a fricative, /v/ is more sonorous than /s/ since voiced fricatives are more sonorous than voiceless fricatives (section 1.2).

acquisition adhere to different sonority-based measurements of complexity, as discussed in section 1.3. Moreover, these measurements are established based on the Sonority Hierarchy, which can be assumed to be universal, or language-specific. The study conducted in this thesis aims to investigate whether different complexity measurements can make different predictions about the effect of sonority on accuracy in producing complex onsets. Therefore, both the MSD and the SDP were obtained, based on the formulas presented in section 1.3. Both measurements of sonority-based markedness were obtained based on the Sonority Hierarchy specific to Icelandic proposed by Heimisdóttir (2014), shown in figure (5), and followed her classification of sounds in treating both /v/ and /j/ as approximants. According to this scale, /s/ have lower sonority than other fricatives, and are classified as the least sonorous together with aspirated stops. The classification of clusters based on the MSD, with initial onsets representing given sonority profiles, and their frequency in the dataset, is presented in table (7). The smaller sonority distance between two consonants in the complex onset indicates a more marked status of this cluster. Thus, *s + stop* sequences are the most marked, while the clusters consisting of *s + approximants* and *aspirated stop + approximant* have the greatest distance on the sonority scale and are the least marked.

Sonority Distance	Sonority profile	Clusters in the dataset	Percentage of the total dataset
1 (most marked)	*s + stop*	**/sp/, /sc/, /sk/, /st/**	22%
2	*lateral + approximant* *fricative + lateral* *nasal + approximant*	**/lj/** **/fl/** **/mj/**	11%
3	*stop + lateral* *fricative + trill*	**/pl/, /kl/** **/fr/, /θr/**	19%
4	*stop + trill* *aspirated stop + lateral* *fricative + approximant* *s + nasal*	**/kr/, /pr/, /tr/** **/pʰl/, /kʰl/** **/θv/** **/sn/, /sm/**	33%
5	*aspirated stop + trill* *stop + approximant*	**/pʰr/, /tʰr/, /kʰr/** **/pj/**	12%
6 (least marked)	*aspirated stop + approximant* *s + approximant*	**/tʰv/** **/sv/, /sj/**	9%

Table 7. The SD in initial clusters based on the Sonority Hierarchy specific to Icelandic.

The classification based on the Sonority Dispersion values of initial clusters is presented in table (8). With the increase of the dispersion value, the markedness of a given onset increases. The least marked complex onsets have the most equally dispersed sonority value between two cluster members and the vowel.

Dispersion value	Sonority profile	Clusters in the dataset	Percentage of the total dataset
1,36 **(most marked)**	*lateral + approximant* *nasal + approximant*	**/lj/** **/mj/**	6%
1,1	*fricative + approximant*	**/θv/**	3%
1,07	*stop + approximant* *s + stop*	**/pj/** **/sp/, /sc/, /sk/, /st/**	25%
1,05	*aspirated stop + approximant* *s + approximant*	**/tʰv/** **/sv/, /sj/**	5%
0,4	*fricative + trill* *fricative + lateral*	**/fr/, /θr/** **/fl/**	11%
0,34	*stop + trill*	**/kr/, /pr/, /tr/**	14%
0,31	*aspirated stop + trill*	**/pʰr/, /tʰr/, /kʰr/**	8%
0,25	*stop + lateral*	**/pl/, /kl/**	14%
0,19 **(least marked)**	*aspirated stop + lateral* *s + nasal*	**/pʰl/, /kʰl/** **/sn/, /sm/**	14%

Table 8. Dispersion values of initial clusters based on the Sonority Hierarchy specific to Icelandic.

Secondly, the measurements were obtained based on the universal Sonority Hierarchy[7], presented in section 2, to investigate whether assuming a different sonority scale leads to different predictions. Here, /s/ is classified with other fricatives rather than being treated as a less sonorous segment. In this classification, /v/ in initial clusters is analyzed as a fricative. Further, /r/ and /l/ are assumed to be equally sonorous (labeled as *liquids*), and there is no

---

[7] I refer to the most cited Sonority Hierarchy as *universal* for the sake of simplicity. However, it must be noted that this does not indicate that this hierarchy is "absolute", and there is a debate on that matter (e.g., Parker 2002).

distinction between aspirated and unaspirated stops. Classification based on Sonority Distance is presented in table (9). The most marked clusters are composed of *s + stop* since these segments comprise sonority reversals. The least marked clusters have the steeper rise in sonority from the first member of a cluster to the second.

Sonority Distance	Sonority profile	Clusters in the dataset	Percentage of the total dataset
-1 **(most marked)**	*s + stop*	**/sp/, /sc/, /sk/, /st/**	22%
0	*s + fricative* *fricative + fricative*	**/sv/** **/θv/**	5%
1	*s + nasal* *liquid + approximant* *fricative + stop*	**/sn/, /sm/** **/lj/** **/tʰv/**	11%
2	*fricative + liquid* *nasal + approximant*	**/fl/, /fr/, /θr/** **/mj/**	14%
3	*stop + liquid* *s + approximant*	**/pʰl/, /pʰr/, /tʰr/, /kʰl/, /kʰr/, /pl/, /pr/, /tr/, /kl/, /kr/** **/sj/**	44%
4 **(least marked)**	*stop + approximant*	**/pj/**	3%

Table 9. The SD in initial clusters based on the universal Sonority Hierarchy.

The SDP does not account for sonority reversals, thus *s + stop* and *fricative + fricative* clusters could not be included in this classification (which eliminates 648 observations). Dispersion values of onset clusters based on the universal Sonority Hierarchy are presented in table (10).

Dispersion value	Sonority profile	Clusters in the dataset	Percentage of the total dataset
2,1 **(The most marked)**	*lateral + approximant*	**/lj/**	4%
1,36	*nasal + lateral*	**/mj/**	4%
1,17	*fricative + approximant* *fricative + nasal*	**/sj/, /sv/** **/sm/, /sn/**	15%
1,1	*stop + fricative* *stop + approximant*	**/tʰv/** **/pj/**	8%

0,56	*fricative + liquid*	**/fr/, /θr/, /fl/**	12%
0,4 **(The least marked)**	*stop + liquid*	**/pʰl/, /pʰr/, /tʰr/, /kʰl/, /kʰr/, /pl/, /pr/, /tr/, /kl/, /kr/**	57%

Table 10. Dispersion values of initial clusters based on the universal Sonority Hierarchy.

3.3.3 The measurement of frequency

To investigate the effect of the frequency of clusters in the ambient language on cluster production, the phonotactic probability of initial onset clusters in Icelandic was obtained using a function of the software Phonological CorpusTools (PCT, Hall et al. 2021). The phonotactic probability of initial clusters was calculated based on Icelandic Pronunciation Dictionary (Nikulásdóttir et al. 2022). The Icelandic pronunciation dictionary contains 65 282 words transcribed in four pronunciation variants. Probabilities of initial onset clusters were obtained based on the word list with standard pronunciation with the algorithm proposed by Vitevitch & Luce (2004), which is implemented in PCT. The algorithm uses bigram positional probabilities, which capture how likely the two segments are to occur at a specific position in a word. For initial clusters, the algorithm calculates how likely a given cluster is to occur at the initial position of a word. Since the probabilities were calculated using the Icelandic Pronunciation Dictionary, the obtained probabilities measure type frequency of initial clusters in Icelandic (figure 7).

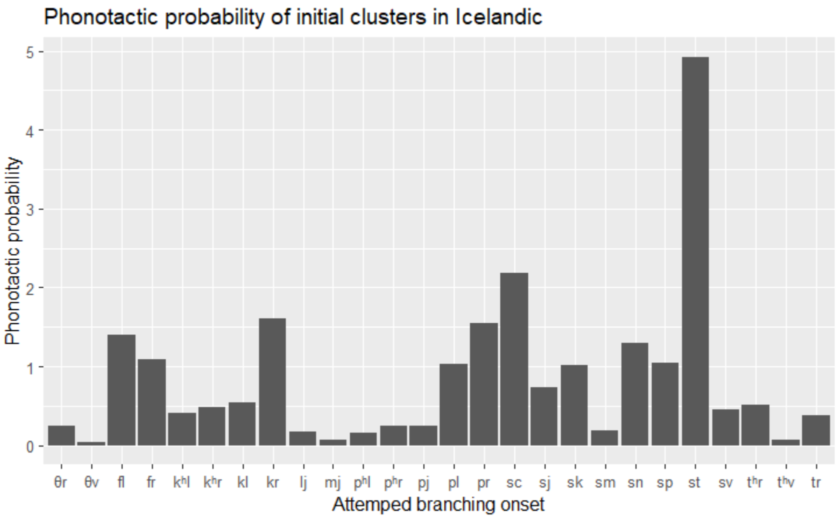


Figure 7. Phonotactic probabilities of initial onset clusters in Icelandic

Unfortunately, no child-directed speech corpora are currently available for Icelandic. There is also no transcribed frequency dictionary that could be used for investigating the effect of token frequency. Orthographical forms do not always correspond to the pronunciation, especially regarding aspirated stops, which sometimes lose aspiration (Heimisdóttir 2015). For instance, orthographic <t> can be used to represented both /t/ (e.g., *sitja* is transcribed as /ˈsɪːtja/) and /tʰ/ (*tromma* is transcribed as /tʰrɔmma/). Thus, only the effect of type frequency of initial clusters on accuracy in producing clusters was investigated.

3.3.4 The measurement of accuracy in producing singleton consonants and the effect of articulatorily difficult segments.

Additionally, I investigated whether the children who mastered more single segments are more accurate at producing clusters. As discussed in section 2.2, in some instances even though a segment was acquired and a child produces it in another position of a word, the cluster is still reduced to one consonant. The accuracy score was obtained for each participant based on accuracy annotations, included in the Másdóttir Corpus, calculated by dividing the number of attempted consonants by the number of correctly produced consonants. The accuracy score for all children is listed in Appendix I.

To investigate whether the order of acquisition of singleton consonants influences the order of acquisition of initial clusters, the group of fricative-initial clusters (1139 observations) was compared to the group of stop-initial clusters (335 observations). Additionally, all clusters were classified based on the manner of articulation of their consonants, to investigate whether the articulatory difficulty of trills affects the accuracy in producing initial onsets with this segment. This classification was also performed to investigate whether the phonemes /j/ and /v/ pattern with other sonorants when these phonemes occur as a second member of a tautosyllabic cluster. Thus, branching onsets and sC clusters were grouped based on their manner of articulation, as presented in tables (11) and (12).

Manner of articulation of singleton consonants	Clusters in the dataset	Frequency among all branching onsets
lateral + approximant / fricative	**/lj/**	4%
nasal + approximant/fricative	**/mj/**	4%
stop + approximant/ fricative	**/tʰv/, /pj/**	8%
fricative + approximant/voiced fricative	**/θv/**	4%
fricative + lateral	**/fl/**	8%

fricative + trill	**/fr/, /θr/**	8%
stop + trill	**/pʰr/, /tʰr/, /kʰr/, /pr/, /tr/, /kr/**	33%
stop + lateral	**pʰl/, /kʰl/, /pl/, /kl/**	29%

Table 11. Classification of branching onsets based on the manner of articulation of individual segments.

Manner of articulation of singleton consonants	Clusters in the dataset	Frequency among all sC clusters
s + nasal	**/sm/, /sn/**	17%
s + approximant/ fricative	**/sv/, /sj/**	17%
s + stop	**/sp/, /st/, /sc/, /sk/**	66%

Table 12. Classification of sC clusters based on the manner of articulation of individual segments

4. Results and analysis

This chapter presents the results of the data analysis. Statistical analysis was conducted in the free statistical software R Studio version 4.2.3. (R Core Team 2021). R Studio was also used to make all graphs and plots presented in this section. Firstly, in subsection 5.1., I present descriptive statistics of the overall performance on different types of initial consonant clusters, and in section 5.2., reduction patterns will be presented. In subsections 5.3., 5.4., and 5.5., the descriptive statistics of investigated factors, i.e., sonority, phonotactic probability, accuracy on singleton consonants, and the natural classes of segments will be presented to determine which factors affect children's accuracy. In subsection 5.6., I present the quantitative analysis of the discussed factors.

4.1 Overall accuracy of onset cluster production

First, the accuracy in producing branching onsets was compared with the accuracy in producing bi- and triconsonantal sC clusters. Figure (8) shows the proportion of accurate responses for each cluster type, which reveals that branching onsets were produced with the highest accuracy. It must be noted that all the figures presented in this subsection and the next, show the relative proportions, rather than the absolute count. The salmon-colored areas show the percentage of inaccurate production, while the turquoise-colored areas show the percentage of accurate productions.

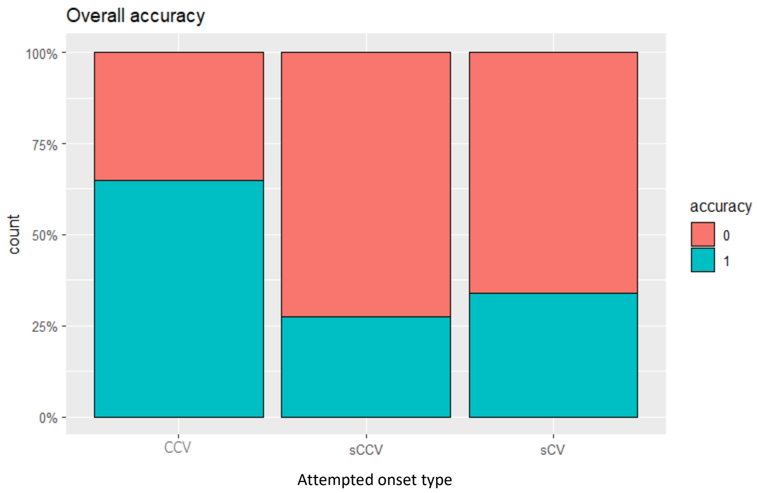


Figure 8. Overall accuracy in producing different onset types.

Within the class of branching onsets, children exhibited the lowest accuracy on clusters containing initial fricatives – /θr/, /θv/, /fl/, and /fr/ – as shown in figure (9).

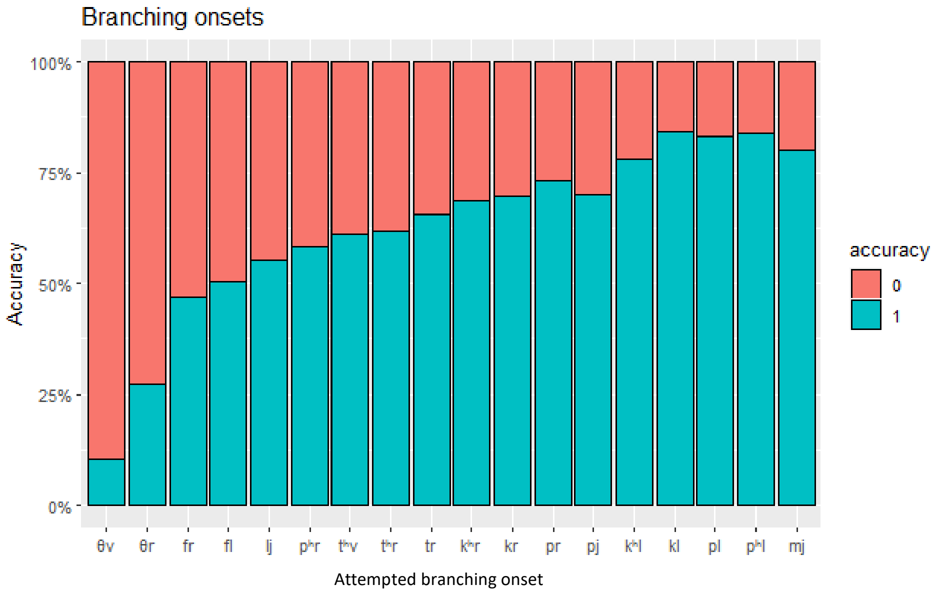


Figure 9. Accuracy in branching onsets.

With regard to sC clusters, children produced /sj/ with the greatest accuracy compared to other sC clusters. As figure (8) reveals, overall accuracy on sC clusters was low in comparison to branching onsets. Moreover, children exhibited equally low accuracy on *s* + *stop* clusters, which comprise sonority reversals, and *s* + *nasal* clusters, which comprise sonority rises. Additionally, children showed lower accuracy on /sv/ cluster than /sj/. This might indicate that /v/ does not pattern with /j/ and should be classified as a fricative. Thus, /sv/ might comprise a sonority plateau, rather than a sonority rise with a sonority distance of 3 like /sj/. Due to this observation, in the investigation of the MSD effect based on the universal Sonority Hierarchy (table 9), /j/ was analyzed as an approximant, and /v/ was analyzed as a fricative.

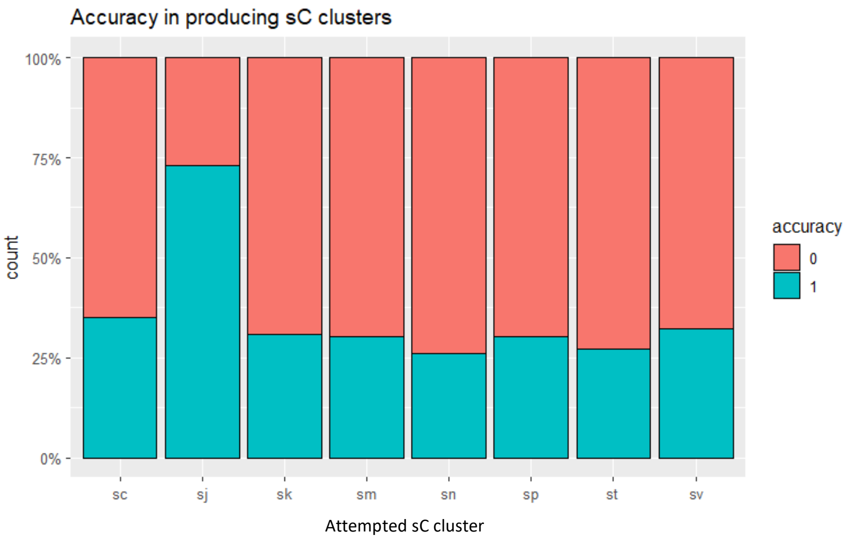


Figure 10. Accuracy in biconsonantal sC clusters

Children showed low accuracy on triconsonantal clusters, as shown in figure (11).

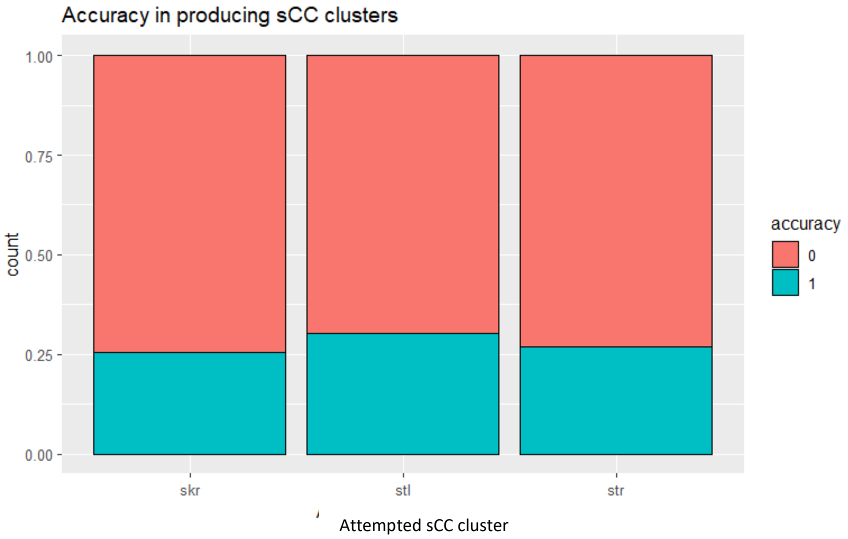


Figure 11. Accuracy in triconsonantal sCC clusters

4.2 Reduction patterns

The investigation of reduction patterns revealed that branching onsets and sC clusters are targeted by different patterns. Moreover, slightly different patterns occur in respective clusters, as shown in figure (12) summarizing the proportions of reduction types targeting branching onsets.

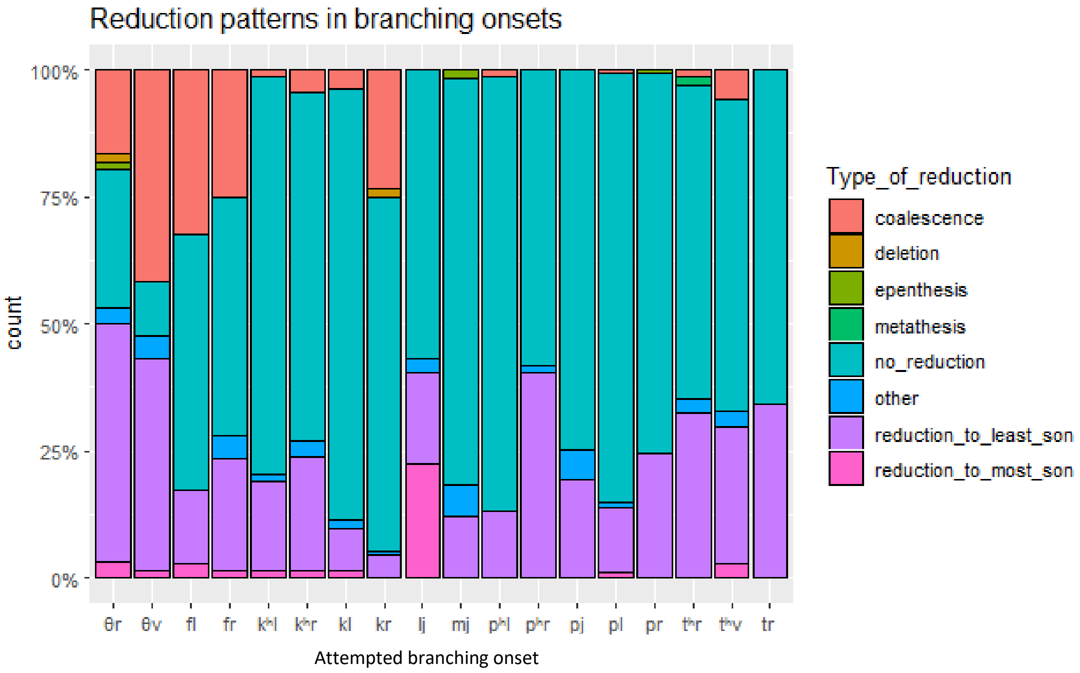


Figure 12. Reduction patterns occurring in branching onsets.

Table (13) summarizes the most and the least common simplification patterns targeting the whole group of branching onsets. The most common simplification pattern was a reduction to the least sonorous consonant in the cluster. This targeted almost all branching onsets, and resulted in the deletion of the second, more sonorous consonant. The second commonly occurring process was coalescence. This process mostly occurred when there was a possibility for two phonemes to merge into one that retained the features of both. Thus, it occurred the most frequently in fricative-initial clusters: /θv/, /θr/, /fl/, /fr/, as well as in the cluster /kr/. Reduction to the most sonorous consonant occurred only in the cluster /lj/.

Simplification process	Frequency among all branching onsets
no reduction (produced as accurate)	66,2%
reduction to the least sonorous	20,4%
coalescence	9,1%
reduction to the most sonorous	1,9%
epenthesis, deletion, metathesis, and other processes	2,37 %

Table 13. Summary of simplification processes in branching onsets

Figure (13) shows the most common simplification processes targeting individual clusters, and table (14) summarizes the most and the least common simplification patterns targeting the whole group of sC clusters .

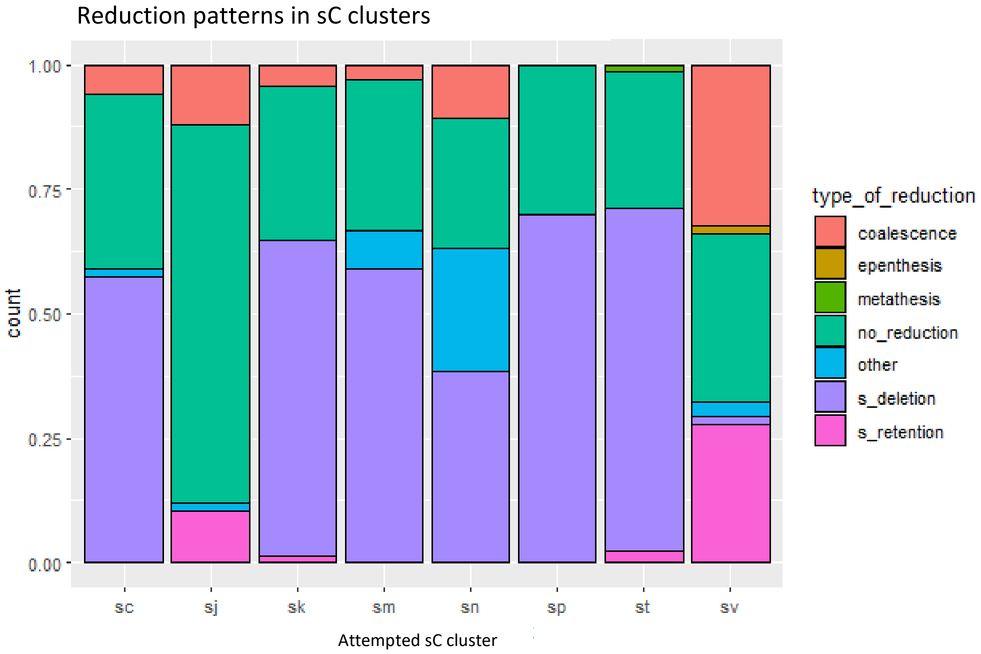


Figure 13. Reduction patterns occurring in sC clusters

In the group of *s* + *stop* clusters, comprising sonority reversals, the most frequently retained was the least sonorous consonant which in this case is the stop, and the initial /s/ was deleted. Rising sonority sC clusters composed of *s* + *nasal* were most frequently reduced to the nasal, which is the most sonorous segment. Thus, the initial /s/ was deleted in either case. Both *s* + *stop* and *s* + *nasal* clusters were sometimes targeted by coalescence. For instance, /sn/ was frequently realized as /n̥/ where the voiceless feature of the initial sibilant was present on the sonorant. Coalescence was the most common process operating in the cluster /sv/. This cluster was also frequently reduced to the least sonorous consonant, which in that case is the initial /s/. The retention of the initial /s/ was also common in /sj/ clusters, which was produced with the greatest overall accuracy. Again, the discrepancy in reduction patterns supported the suggestion that /j/ and /v/ do not bear equal sonority values and do not share the manner of articulation.

Simplification process	Frequency among all sC clusters
no reduction (produced as accurate)	34,7%
/s/ deletion	51,3%
/s/ retention	3,6%
coalescence	6,6%
epenthesis, deletion, metathesis, *and other processes*	3,7 %

Table 14. Summary of simplification processes in sC clusters

The deletion of the initial /s/ also occurred commonly in all triconsonantal onsets (table 15, figure 14). Moreover, after the initial /s/ was deleted, the other two consonants were often reduced to the least sonorous one, i.e., /str/ in /straiːtou/ was reduced to [t] in [taiːtou]. Reduction to the least sonorous consonant combined with retention of the initial /s/, as well as other simplification processes occurred rarely in this cluster type.

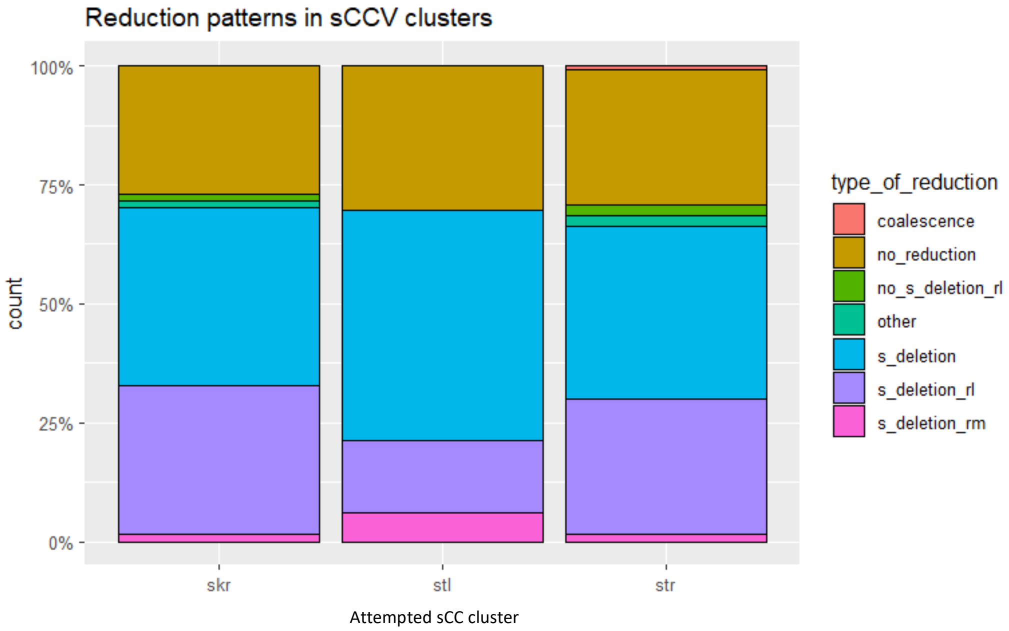


Figure 14. Reduction patterns occurring in sCC clusters.

Simplification process	Frequency among all sC clusters
no reduction (produced as accurate)	28,5%
/s/ deletion	39,7%
/s/ deletion and reduction to the least sonorous segment	25,8%
/s/ deletion and reduction to the most sonorous segment	2,6%
/s/ retention and reduction to the least sonorous segment	3,7 %
coalescence and other processes	1,8%

Table 15. Summary of simplification processes in sCC clusters

4.3 Sonority-based markedness

The MSD and the SDP were calculated based on the Sonority Hierarchy specific to Icelandic. The clusters were classified based on their sonority profiles, as demonstrated in table (7) and table (8). Figures (15) and (16) demonstrate that contrary to predictions, the complexity measurements calculated using the Icelandic-specific Sonority Hierarchy, children accuracy did not decrease gradually according to the complexity status of initial clusters. The MSD predicts that clusters with the smallest increase in sonority between the first and the second consonant are produced with the least accuracy, and that clusters with the greater increase in sonority are produced with the greatest accuracy. Children did exhibit the lowest accuracy in producing the clusters with the smallest sonority distance (henceforth SD), but their accuracy did not increase with the increase of SD between cluster members.

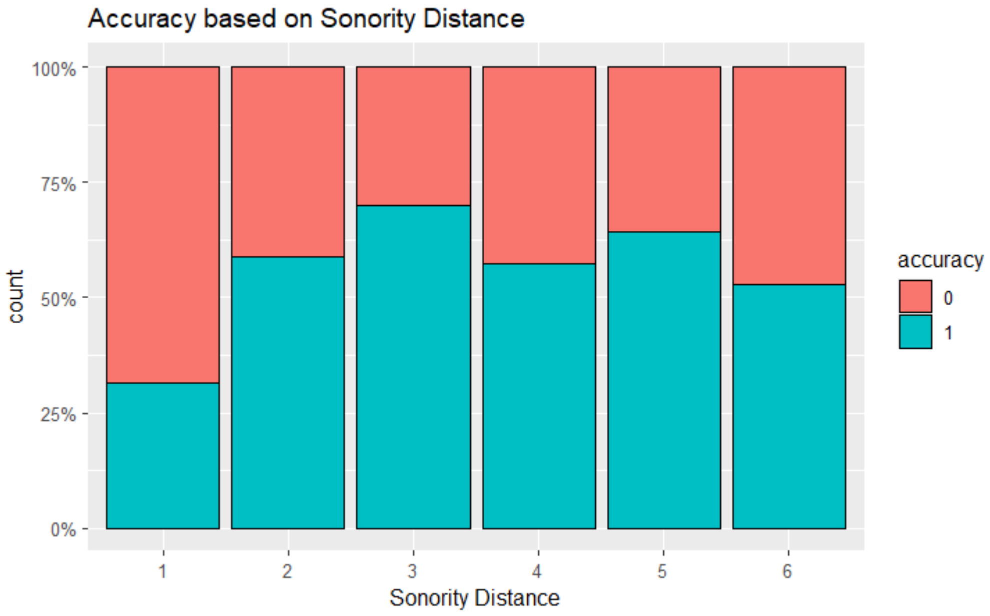


Figure 15. Association between accuracy on initial clusters and SD, based upon the Sonority Hierarchy specific to Icelandic.

The cluster complexity based on the SDP also does not correlate with a gradual increase in accuracy. The children did not exhibit the lowest accuracy on the most marked clusters with the greatest dispersion value. Thus, neither the MSD nor the SDP could predict the tendencies in children's accuracy in producing onsets clusters.

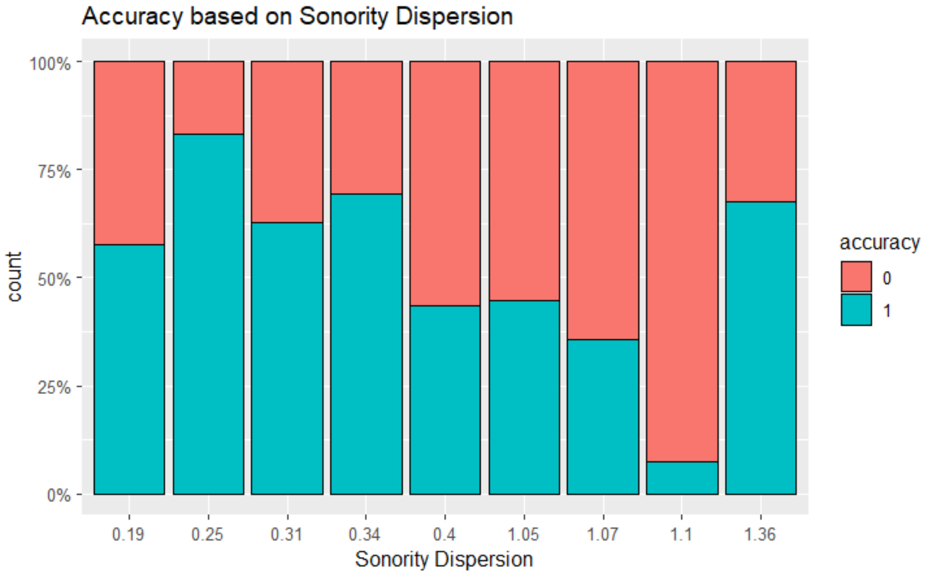


Figure 16. Association between accuracy on initial clusters and the dispersion values, based upon the Sonority Hierarchy specific to Icelandic

Likewise, the complexity ranking based on the dispersion values calculated based on the universal Sonority Hierarchy did correlate with accuracy in producing onsets (figure 17). In this case, the onsets with the lowest dispersion value were produced with the greatest accuracy, but those with the dispersion value of 1.36 were produced with equal accuracy. Furthermore, the children showed low accuracy on clusters with a dispersion value of 0.56, and great accuracy on clusters with bigger dispersion values.

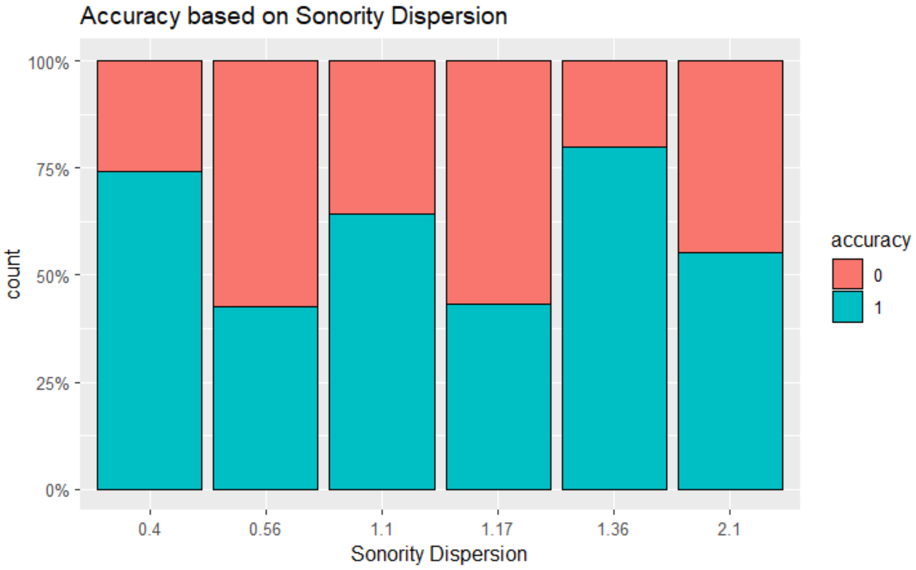


Figure 17. Association between accuracy on initial clusters and the dispersion values based upon the universal Sonority Hierarchy.

The markedness of onset clusters based on the MSD and the universal Sonority Hierarchy seems to make the most correct predictions, as shown in figure (18). Children exhibited the lowest accuracy on sonority reversals and plateaus with the SD of -1 and 0. The accuracy in producing initial onsets increased with the increase of SD. Nevertheless, according to the predictions of the MSD, sonority reversals are more marked than sonority plateaus and children showed the lowest accuracy on sonority plateaus – clusters composed of two fricatives (/sv/, /θv/). As mentioned earlier, in this classification /v/ was treated as a fricative, rather than an approximant. This classification was made to account for the fact that, when these two clusters were treated as *fricative* + *approximant* (the SD value of 3), the analysis exhibited that this cluster was a strong outlier. Children produced this cluster with the lowest accuracy, while *stop* + *liquid* clusters were one of the most accurately produced clusters.

Children produced clusters with the SD of 3, i.e., clusters composed of a stop followed by a liquid, more accurately than those with the SD of 4. However, it must be noted that this class was represented by only one cluster, as shown in table (10). Furthermore, according to the MSD, *s+nasal* clusters with one steep rise in sonority should be produced with greater accuracy than *s+stop* clusters. However, as shown in figure (10), all these clusters are produced with equally low accuracy. This indicates that MSD does not affect these two classes of initial clusters, which will be further explored through quantitative analysis.

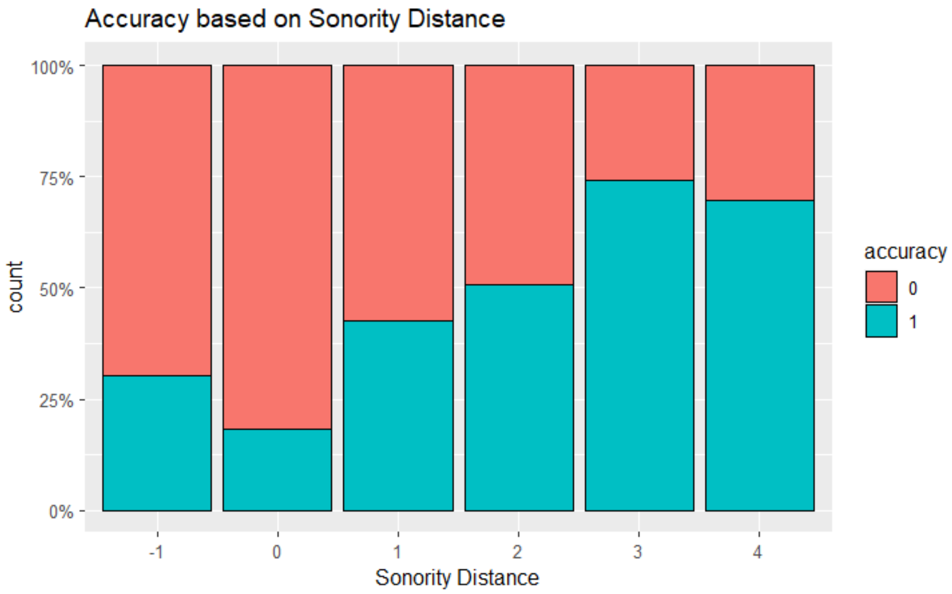


Figure 18. Association between accuracy on initial clusters and MSD based upon the universal Sonority Hierarchy.

As shown in subsection 4.2.4., assuming different Sonority Hierarchies and different phoneme classifications drastically changes the classification of the clusters across the categories of the MSD and the SDP. The SDP and the MSD calculated based on the Sonority Hierarchy specific to Icelandic could not account for the analyzed data, and children's performance did not decrease with the markedness of clusters. Most importantly, as shown in tables (7) and (8), classification based on the Sonority Hierarchy specific to Icelandic groups clusters in which children showed varying accuracy levels into one category. According to the MSD, the most marked clusters are *s + stop,* in which children did show low accuracy. Comparatively, the clusters with the SD of 2 are /lj/, /fl/ and /mj/. Children did show similar low accuracy in the production /lj/ and /fl/, but /mj/ was one of the most accurately produced clusters. The SD of 3 groups *stop + lateral* clusters (/pl/, /kl/) with *fricative + trill* (/fr/, /θr/). clusters, even though these groups vary significantly in the accuracy of their production. The group of SD 4 groups together *stop + trills* and *lateral* with the cluster /θv/ that children produced with the lowest accuracy. Finally, the groups with the most unmarked SD of 5 and 6 groups included *aspirated* *stop + trill* and *obstruent + approximant* clusters. According to Icelandic-specific Sonority Hierarchy, /r/ is more sonorous than /l/. Thus, fricatives and stops followed by a /l/ should be acquired before clusters with /r/ as a second member, which was not the case for the analyzed dataset (figure 9). Moreover, the MSD also favors *aspirated stop + liquid*. However, as shown in figure 9, children were slightly more accurate on clusters with unaspirated stops, even though *aspirated stop + liquid* clusters have a bigger step in sonority.

On the other hand, the SDP favors *stop + lateral* over *stop + trill*. However, the clusters with aspirated stops are still favored over the clusters with unaspirated stops. Most importantly, according to this classification the group of clusters with fricatives with the SDP of 0,4 is favored over the clusters /t^{h}v/, /pj/, /lj/, /mj/, /sj/ which children produced more accurately. For these reasons, I conclude that the predictions of the MSD and the SDP based on the fine-grained Sonority Hierarchy specific to Icelandic could not account for the patterns observed in the data, and for this reason I do not investigate them further in the quantitative analysis.[8]

The SDP that was calculated based on the more coarsely grained Sonority Hierarchy favors *stop + liquid* clusters, which were the most accurately produced by children. However, it also favors the group of *fricative + liquid* clusters, although these were produced with the lowest accuracy in the clusters /t^{h}v/, /pj/, /lj/, /mj/, and /sj/. Nonetheless, the MSD based on the universal Sonority Hierarchy seems to make the most correct predictions. Figure (18) shows a trend towards accuracy with the increase of sonority distance. Hence, to examine whether the effect of the MSD is statistically significant, its association with children's accuracy is investigated further through quantitative analysis in the next section.

4.4 Phonotactic probability

There was no association between accuracy and the phonotactic probability of initial onset clusters, as shown in figure (19). Thus, children did not show greater accuracy in producing clusters which frequently occur in Icelandic. In fact, high probability sC were produced with the lowest accuracy, as presented in figure (7).

---

[8] The quantitative analysis confirmed that both measurements based on Sonority Hierarchy specific to Icelandic either did not show a positive correlation between less marked clusters and children's accuracy, or any significant associations, as observed in figures (15-17). It must also be noted that the models were fitted to the datasets with, and without sC clusters. For reasons of spatial limitation, I do not show and discuss these models.

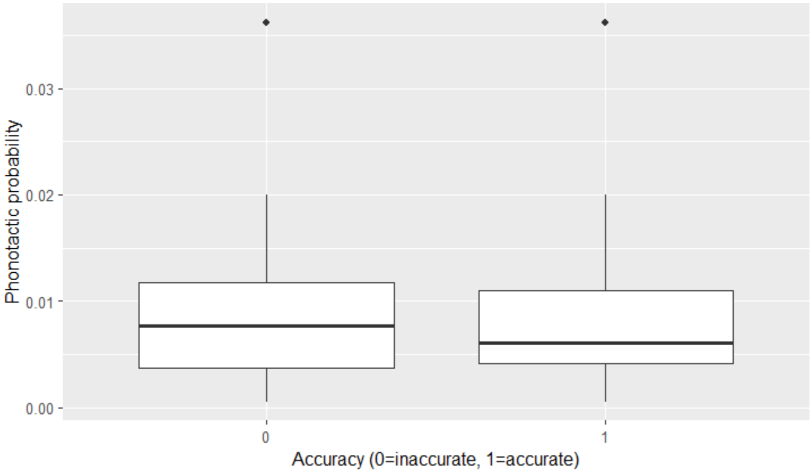


Figure 19. Association between accuracy and phonotactic probability of initial onset clusters.

4.5 Accuracy on singleton consonants

Higher accuracy in producing singleton onsets was associated with higher accuracy on consonant clusters, as revealed by figure (20). The plot on the right side shows that children whose accuracy score was centered around 80% produced initial onsets most accurately, while the plot on the left side indicates that subjects with an accuracy score centered around 65% produced clusters least accurately. The upper and lower whiskers of both box plots show the great variation in inaccurate and accurate cluster production. Thus, despite their accuracy score in the production of singleton consonants, children exhibited variation in producing initial clusters.

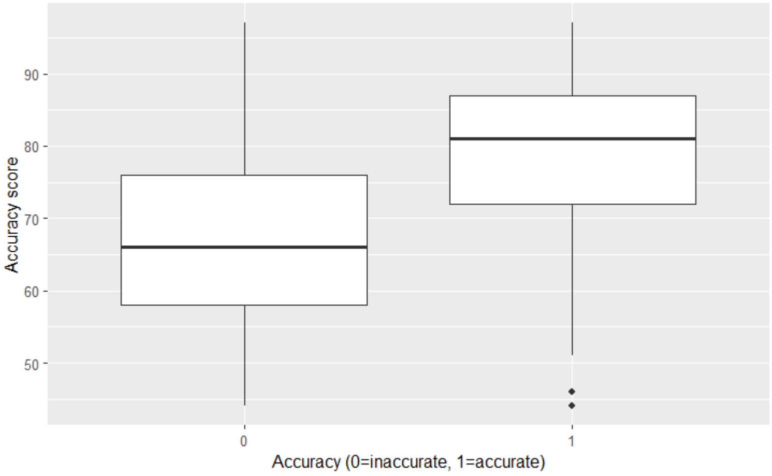


Figure 20. Association between accuracy in initial onsets and singleton consonants.

4.6 The effect of articulatory complexity of individual segments

To answer the question of whether stop-initial clusters are acquired before fricative-initial clusters, the accuracy of clusters constructed of *fricative* + *approximant* was compared to the *stop* + *approximant* clusters (1139 observations of *stop* + *approximant* and 335 observations of *fricative* + *approximant*). Figure (22) demonstrates that fricatives followed by approximants were produced more often as a singleton consonant than stops followed by approximants. To confirm whether these observations are statistically significant, the correlation between accuracy and the natural classes of clusters is further explored through the quantitative analysis presented in the following subsection.

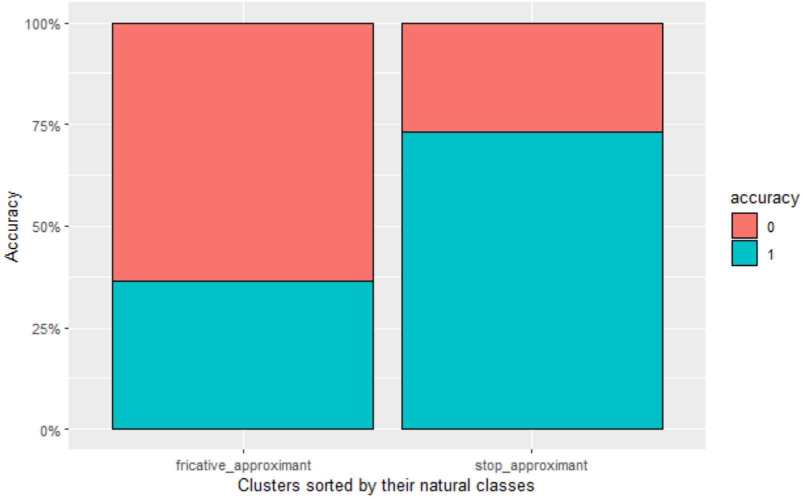


Figure 21. Association between accuracy and stop- versus fricative-initial clusters.[9]

Additionally, all clusters, categorized by their manner of articulation, were sorted by observed accuracy (figure 22). This comparison was performed to investigate whether the phonetic difficulty of other singleton consonants can affect the accuracy in clusters that include these singletons. In addition, accuracy scores were compared to assess whether /v/ and /j/ pattern with other sonorants as a second member of a cluster.

In branching onsets, children showed the lowest accuracy on clusters with a fricative as a first member (figure 21). The results indicated that phonemes /v/ and /j/ when occurring as a second member of a cluster, seem to behave more like approximants than fricatives. Children produced

---

[9] The label approximant groups lateral, trills, and approximants.

the group of clusters *stop* + /v/, /j/ (/tv/, /pj/) with an accuracy equal to *stop* + *trill* clusters. The cluster /mj/ was produced with a level of accuracy almost as high as *stop* + *lateral* clusters. The cluster that deviates from this pattern is /θv/, with the lowest accuracy in production. Children produced this cluster most frequently as a single phoneme. As presented in section 5.2, coalescence was the most frequent reduction pattern that targeted this cluster. Coalescence occurred because these two phonemes share the [+continuant] feature and can be merged into one phoneme that shares the features of two consonants. Additionally, figure (22) demonstrates that the children produced fricatives and stops followed by a lateral with higher accuracy than when followed by a trill.

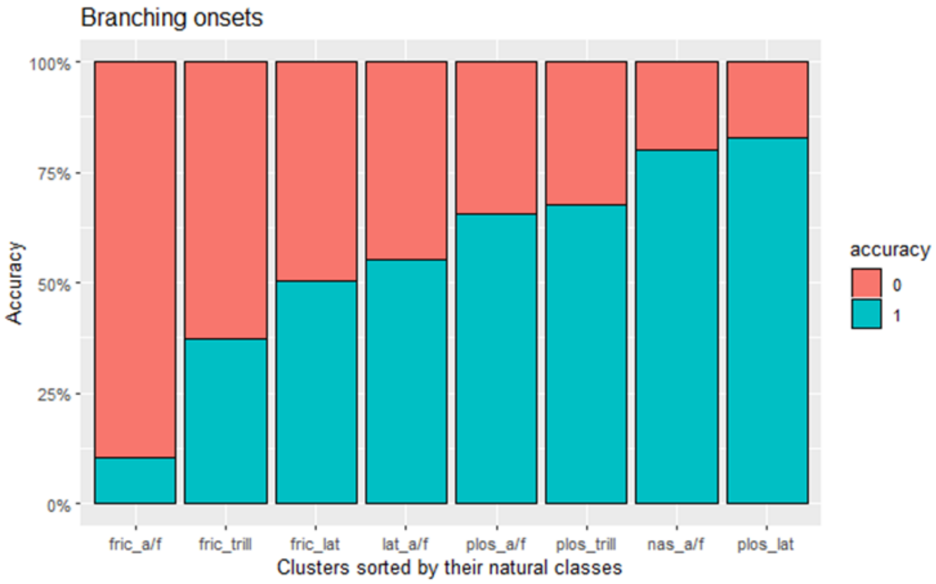


Figure 22. Association between accuracy in branching onsets categorized by their manner of articulation matched by the observed accuracy[10]

Comparison of accuracy of sC clusters revealed that the *s+nasal* clusters were produced with an accuracy level equal to that of *s+stop*. However, there was a discrepancy between the accuracy in production of /sj/ and /sv/. From one perspective, this may indicate that /j/ and /v/ are phonemes with different features, and that /v/ should be treated as fricative while /j/ is treated as a sonorant, since accuracy on /sj/ matches accuracy on other clusters that have a sonorant as a second member. However, as in the cluster /θv/, both consonants in /sv/ share a manner feature; thus, coalescence is possible. Indeed, the /sv/ cluster was the most common

---

[10] *fric* – fricatives, *a/f* – approximants/fricatives, *lat* – lateral, *plos* – plosive stops, *nas* – nasals

target of coalescence among all sC clusters, as shown in figure (12). Therefore, the observed difference may be caused by the possibility of merging two consonants into one segment.

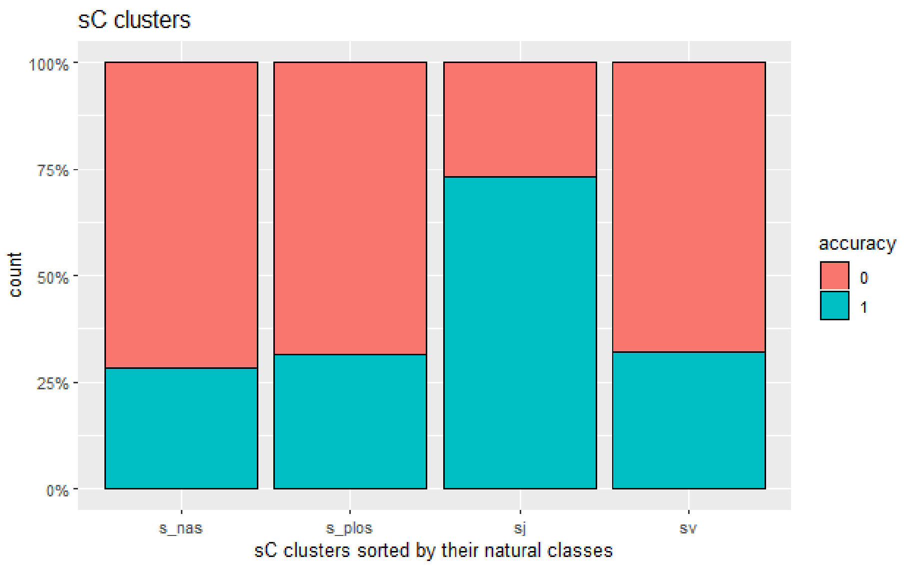


Figure 23. Association between accuracy in sC clusters and natural classes

4.7 Quantitative analysis – modeling the data

This analysis investigates how the discussed factors influence children's accuracy on consonant clusters, and which factors make the best predictions about the process of acquisition. To answer the research questions, I analyzed the correlation between the binary response variable and the independent variables: two measurements of sonority-based markedness (the SDP and the MSD), the accuracy of children on singleton segments, natural classes of clusters, and phonotactic probability of initial consonant clusters in Icelandic. The binary response variable indicated whether a given consonant cluster was produced accurately or inaccurately. Additionally, to control for the effect of potential confounding variables, the age (measured in months) and gender of the children were included as predictors.

Considering all the given variables and the fact that the number of observations of each cluster is not balanced, I selected mixed-effect binomial logistic regression to determine which predictor variable is significantly correlated with children's accuracy on onset clusters. The mixed-effect logistic regression allows for modeling the relationship between several categorical and numerical predictors with the binary response variable, and simultaneously incorporates random effects. The analyzed dataset includes repeated measurements as it includes several observations from each child. Also, all children were asked to pronounce the same set of words. Therefore, the analyzed data does not meet the assumption of independence of models without random effects (Gries 2013). Accordingly, random intercepts were included for the tested items and children's IDs, allowing the model intercept to be adjusted for each individual and item to avoid correlations between data points (Winter 2019, p. 232)[11].

The model was fitted to the dataset using the glmer function from the R package lme4 (Bates et al. 2015). The binary response variable was coded as 1 for accurately produced clusters and 0 for inaccurately produced clusters under the classification presented in table (6). The distinct independent variables of *age*, *phonotactic probability* and *segment accuracy* were entered as numeric predictors, while *gender* and *natural classes* were entered as factor variables (according to the classifications presented in tables (7 - 12)). Also, all clusters in the dataset were coded for *onset type* to distinguish between branching onsets and sC clusters. To investigate whether the MSD, the SPD, or *natural classes* of clusters makes the best predictions, each of the measurements – described in subsection 4.2.4.-5. – were added one after another as independent variables. To compare whether the *sonority distance* (the MSD) or *sonority dispersion* (the SPD) makes the best predictions, a model with these two variables as numeric predictors was fitted to the dataset without sonority reversals and plateaus, for which the SDP cannot account for (1474 observations). The results showed that the effect of SDP was not found to be significant ($\beta = 0.479$, $z = 8.28$, $p = 0.407$), which confirmed the observations shown in figures (17) and (18). The factor *sonority distance* positively correlated with the accuracy ($\beta = 1.511$, $z = 5.31$, $p < 0.001$), Thus, *sonority distance* was entered as a factor variable to investigate its effect more thoroughly.[12]

The effect of MSD combined with the effects of other predictors was investigated, and the analysis included the whole dataset containing both branching onsets and sC clusters. The analysis confirmed that phonotactic probability is not correlated with children's accuracy (figure 19). Furthermore, the analysis showed that gender did not have any significant association with accuracy. Age was positively associated with accuracy. However, this correlation was not found to be significant, most likely because the dataset included only

---

[11] The model with random slopes was also fitted to the dataset but did not converge due to the unbalanced dataset. Therefore, the final model includes only random intercepts.

[12] It might be debatable whether *sonority distance* should be treated as a categorical, or a numerical factor. In the study conducted by Jarosz (2017), sonority is treated as a numerical predictor. Treating sonority as numerical variable can indeed show its overall effect, but does not allow to access whether there are significant differences between different levels of this predictor. Therefore, I analyzed it as a distinct, categorical variable.

children who systematically reduced clusters. The model comparison confirmed that *age*, *gender* and *phonotactic probability* were not significant factors, thus, the final model did not include these variables. The best-fitting model included *sonority distance* (table 9) and *segment accuracy* as predictors (figure 24). The detailed outputs of the regression analysis, with tests for multicollinearity and overdispersion for each model with multiple categorical predictor variables, are included in Appendix 2. In all plots presented in this subsection, the log odds were converted into probabilities predicted by the models to simplify the interpretation of the coefficients for the reader.

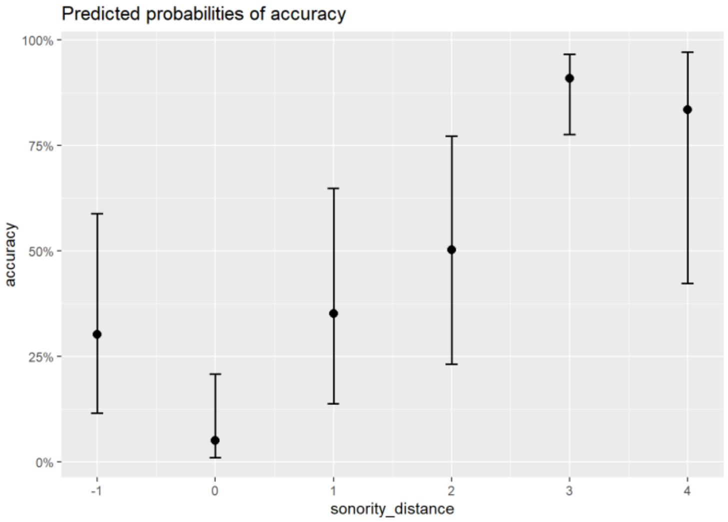


Figure 24. Predicted probabilities of accuracy based on the SD in all clusters.

Since *sonority distance* has 6 levels, to test its main effect the model without sonority distance was compared to the full model. The likelihood ratio test confirmed that the effect of *sonority distance* was significant ($\chi 2(1) = 71.62$, $p < 0.001$). The model revealed a positive correlation between segment accuracy and accuracy in producing both segments in a cluster ($\beta = 0.035$, $z = 2.56$, $p = 0.01$). This result confirmed that the children who acquired more singleton consonants also acquired more initial clusters (figure 20).

The model estimates for clusters with the SD of 0 are significantly lower ($\beta = –1.571$, $z = –2.19$, $p = 0.02$) than for the group of clusters with the SD of -1 (the reference level), i.e., the accuracy for the level of the SD of 0 was lower than for the level of the SD -1. Thus, the analyzed data did not confirm the assumption of the MSD which states that sonority plateaus are acquired before sonority reversals. The model also showed that children demonstrated the greatest accuracy on clusters with the SD of 3 ($\beta = 3.557$, $z = 9.26$, $p < 0.001$), rather than the SD of 4 ($\beta = 3.018$, $z = 3.28$, $p = 0.001$). However, as previously noted, the *sonority distance* level of 4 included only one cluster. The pairwise comparison of all levels of the MSD revealed that these two levels were not significantly different (see Appendix 2). As it can be observed in figure (17), the model also showed that sC clusters which constitute sonority reversals (with the SD of -1, which was the reference level in the presented model), were not significantly different from the clusters with the SD of 1 (/sn/, /sm/, /lj/) (β = 0.652, z= 1.34, p = 0.181). As figure (10) disclosed, the children showed similar accuracy in *s + stop* and *s + nasal* clusters. Therefore, the model validates the observation that accuracy in producing these sC clusters was not affected by the SD. As observed in figure (8), the children produced all sC clusters with lower accuracy than branching onsets. To confirm this observation, the model was run with the predictor variable *onset type*. The results showed that the children produced sC clusters with significantly lower accuracy than branching onsets (β = –2.623 , z= –5.46, p <0.001), and there was a significant effect of *onset type* ($\chi 2(1) = 24.89$, $p < 0.001$). Based on this result and the interpretation of the first model, the model was fitted to the dataset without the sC clusters (1605 observations). Figure (25) presents the association between branching onsets and the level the accuracy, after excluding all sC clusters from the dataset.

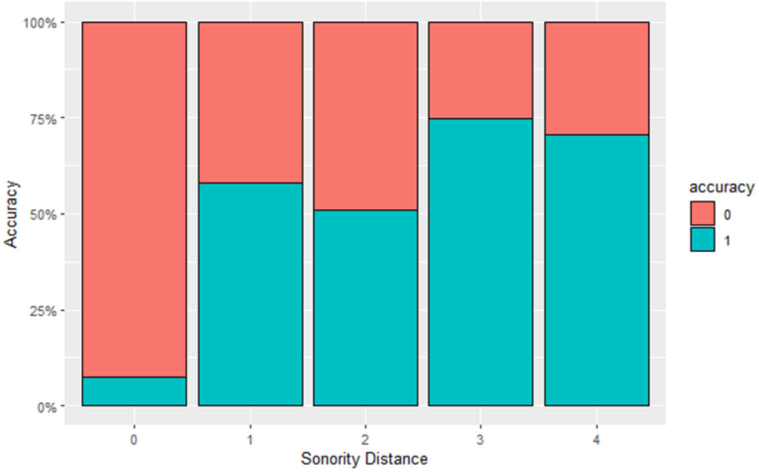


Figure 25. Association between accuracy on branching onsets and the SD

The ANOVA comparison against the null model confirmed the significant effect of *sonority distance* ($\chi 2(1) = 37.22$, $p < 0.001$). Figure (26) reveals the predicted probabilities based on the SD of initial onsets. The reference level in this model was represented by the group of clusters with the SD of 1. After excluding sC clusters, the model captured the highest accuracy in clusters with the SD of 3 (β = 1.618, z = 2.59, p = 0.0096) and the lowest accuracy in clusters with SD of 0 (β = –4.923, z = – 4.25, p < 0.001). Clusters with the SD of 2 were produced with lower accuracy than clusters with the SD of 1, as shown in figure (25). Also, according to the model’s predictions, the group of clusters with the SD of 1 has a higher probability of being produced as accurately than the clusters with the SD of 2 ($\beta = -0.911$, $z = -1.31$, $p = 0.191$) (figure 26). Nevertheless, the pairwise comparison revealed that these, and several pairings of the sonority distance levels were not significantly different (see Appendix 2). Hence, the model showed that the clusters with the SD of 1 had almost the same probability of being produced as accurately as the clusters with the SD of 2. Thus, even though the classification of clusters according to MSD makes correct predictions about accuracy on the most unmarked clusters with the SD of 3 and 4, and predicted the lowest accuracy on the clusters with the SD of 1, the analysis did not confirm all the assumptions of this generalization.

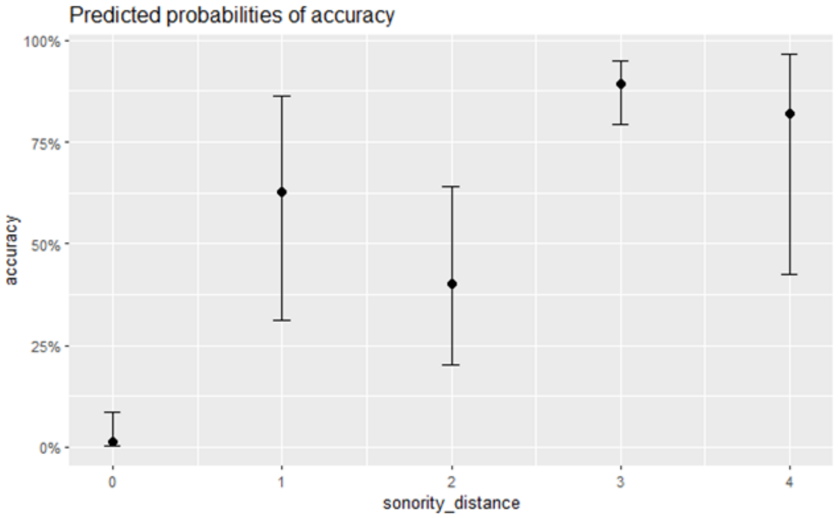


Figure 26. Predicted probabilities of accuracy based on the SD in branching onsets.

To assess whether a model with *segment accuracy* or *sonority distance* fits the data better and makes the best predictions, a nested model comparison with the ANOVA function was fitted. The full model with both predictors was compared against models with *segment accuracy* and *sonority distance.* Additionally, the Akaike Information Criterion (AIC) was compared to assess the goodness of fit of the model for the analyzed dataset[13]. Thus, the results of the nested model comparison showed that the full model fits the data better than a model with one predictor, which indicates that both predictors explain the observed associations. However, the model with *segment accuracy* fits the data better ($\chi 2(1) = 65.14$, $p < 0.001$) than the model with *sonority distance* ($\chi 2(1) = 31.71$, $p < 0.001$).

---

[13] Lower AIC indicates a better fit (Levshina 2015).

	The baseline model	*segment accuracy* + *MSD*	*segment accuracy*	*MSD*
AIC	1236.4	1190.6	1173.3	1212.7
$\chi 2$**(1)**		98.92	65.14	31.71
p		<0.0001	<0.0001	<0.0001

Table 16. The results of nested model comparison for MSD and segment accuracy.

The model was also fitted to the dataset with clusters sorted based on their manner of articulation (*natural classes*) to investigate whether the phonetic difficulty of singleton consonants can influence the accuracy in clusters containing these consonants. As figure (20) revealed, stop-initial clusters were produced as two segments more frequently than fricative-initial clusters. Therefore, at first, a model with *natural classes* and *segment accuracy* as predictors was fitted to the dataset containing these two types of clusters (1474 observations in total). The results confirmed that the main effect of *natural classes* was significant ($\chi 2(1) = 22.40$, $p < 0.001$), and children produced *fricative* + *approximant* clusters with significantly lower accuracy in comparison to clusters composed of *stop* + *approximant* ($\beta = -3.316$, $z = -5.55$, $p < 0.001$). Figure 27 visualizes the model predictions based on these two natural classes of segments (here *approximant* groups together approximants, laterals, and trills).

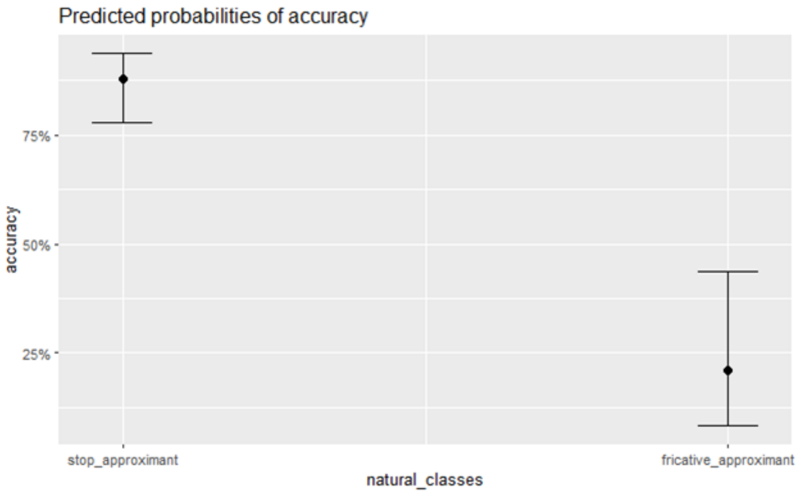


Figure 27. Predicted probabilities of accuracy by natural classes of initial clusters.

Lastly, a more fine-grained distinction between natural classes was investigated to assess the significance of the observation shown in figure (19). Because of too few data points for a level of a predictor variable, three levels were excluded from the model: *fricative* + *approximant*, *lateral* + *approximant*, and *nasal* + *approximant* (as table 11 shows, all these levels construct only 4% of the dataset respectively). Accordingly, the model was fitted to the dataset including 1406 observations. There was a significant main effect of the predictor *natural classes* ($\chi2(1) = 48.96$, $p < 0.001$). The coefficients of the model output showed that the levels *stop* + *trill* ($\beta = -2.096$, $z = -6.28$, $p < 0.001$), *stop* + *approximant* ($\beta = -2.389$, $z = -5.07$, $p < 0.001$), and *fricative* + *lateral* ($\beta = -3.583$, $z = -7.94$, $p < 0.001$), and *fricative* +*trill* clusters ($\beta = -4.719$, $z = -9.6$, $p < 0.001$) were more likely to be produced with lower accuracy with the reference to *stop* + *lateral* clusters. Consequently, the predicted probabilities based on this model indicated that *stop* + *lateral* clusters are likely to be produced accurately as two segments (figure 29). Again, *segment accuracy* positively correlates with accuracy on initial clusters ($\beta = 0.193$, $z = 8.92$, $p < 0.001$).

The pairwise comparison of all the levels of *natural classes* confirmed the observation that *stop* + *lateral* clusters are acquired first, followed by *stop* + *trill* and fricative-initial clusters since there was a significant difference between these two groups. It must be noted that the pairwise comparison revealed no significant difference between the *fricative* + *trill* and *fricative* + *lateral* groups. Additionally, the group of *stop* + *approximant* was not significantly different from *fricative* + *lateral* and *stop* + *trill* (see Appendix 2). As shown in figure (9), the cluster/tʰv/ had lower accuracy than the cluster /pj/. Its accuracy is similar to that of *stop* + *trill* clusters. On the other hand, the cluster /pj/ seems to pattern with the group of *stop* + *lateral* by degree of accuracy. Figures (28) and (29) visualize the model's predictions based on the natural classes of segments and the accuracy in producing singleton consonants.

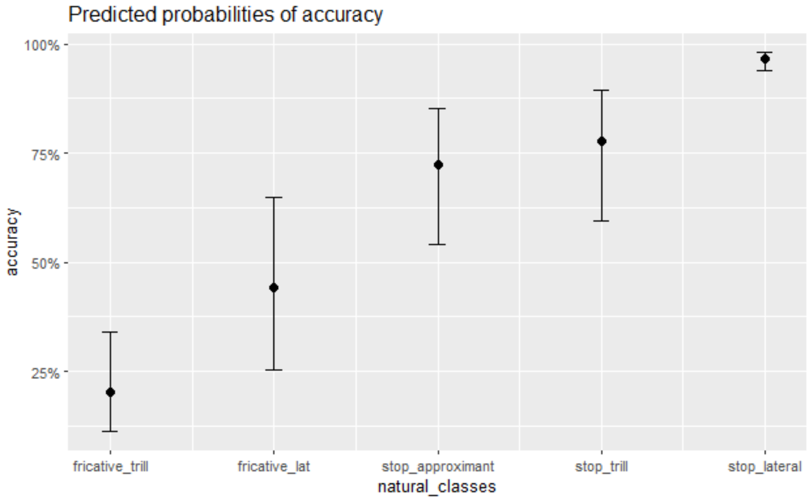


Figure 28. Predicted probabilities of accuracy by fine-grained distinction of natural classes of initial clusters

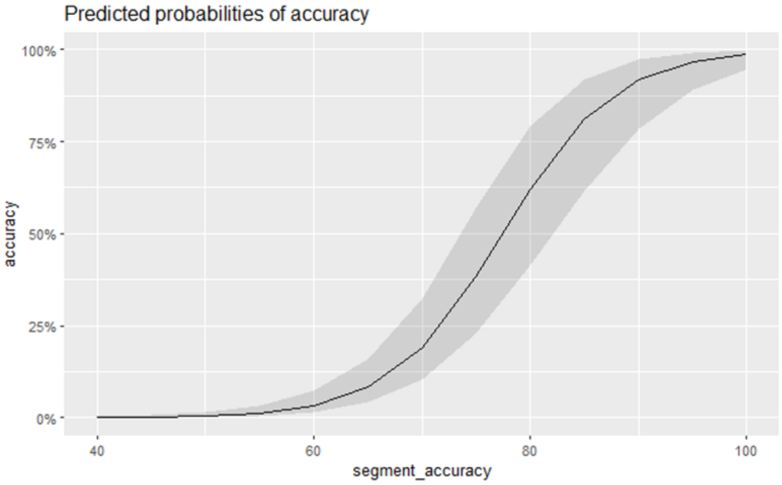


Figure 29. Predicted probabilities of accuracy by the accuracy score on singleton consonants

To assess whether a model with *segment accuracy* or *natural classes* fits the data better and makes the best predictions, a nested model comparison with the ANOVA function. The full model with both predictors was compared to a model with *segment accuracy* and *natural classes*, respectively. Table (14) summarizes the results of the model comparison. The $\chi 2$ value shows the improvement in fit with the comparison to the minimal baseline model, and the *p* value shows the significance of the model comparison under the assumption that the models are the same (Winter 2019, p. 261). The likelihood ratio nested model comparison revealed that the model with both predictors has the best fit ($\chi 2(1) = 116.25$, $p < 0.001$). The results also showed that the model with *segment accuracy* fits the dataset better ($\chi 2(1) = 67.76$, $p < 0.001$) than the model with *natural classes* ($\chi 2(1) = 48.96$, $p < 0.001$).

	The baseline model	*natural classes + segment accuracy*	*segment accuracy*	*natural classes*
AIC	1050.5	944.24	985.2	1010.0
$\chi 2$(1)		116.25	67.76	48.96
p		<0.0001	<0.0001	<0.0001

Table 17. The results of the nested model comparison for predictors *natural classes* and *segment accuracy*

To summarize this subsection, the model with *sonority distance* as a predictor captured the observed low accuracy on the clusters with the SD of 1, and the high accuracy on the group of the clusters with the SD of 3 and 4. According to the MSD, stops followed by the most sonorous approximants are the most unmarked clusters in the dataset. Children showed great accuracy on the cluster /pj/ but produced the clusters /t^{h}v/ and /θv/ with lower accuracy. As mentioned earlier, when the MSD was investigated, /v/ was classified as a fricative. Thus, the cluster /θv/ was analyzed as a sonority plateau with the MSD of 0. Otherwise, if /v/ was treated as approximant, the clusters /θv/ and /sv/ would have been grouped with other clusters with the MSD of 3. However, children showed the lowest accuracy in these clusters compared to other clusters with the MSD of 3, such as *stop + liquid*, which were produced with the highest accuracy. Therefore, if /v/ was treated as an approximant, both clusters were outliers and could not be included in the statistical model.

After excluding the sC clusters and analyzing the effect of the SD on the branching onsets in isolation, the MSD could not predict low accuracy for the group with the SD of 2 that groups *fricative + liquid* clusters, on which children showed the lowest accuracy. Also, the difference between the groups of the SD 1 and 2 was found not to be significant. Thus, both groups had the same probability of being produced as accurate. Generally, any sonority-based generalization does not predict the low accuracy on *fricative + liquid* clusters. Also, the MSD generalizations, based on the universal Sonority Hierarchy, groups clusters with *stop + trill* and *stop + lateral* in the category of the SD 3. If one assumed a more fine-grained Sonority Hierarchy, clusters composed of *stop + trill* have a greater rise in sonority, and children should therefore acquire them before *stop + lateral* clusters, which have a smaller rise in sonority. This assumption goes against the finding that *stop + lateral* clusters were produced with a higher accuracy than *stop +trill clusters* (significance was confirmed by the analysis).

The model with *natural classes* as a predictor showed how grouping segments based on their manner of articulation can predict the increase in the accuracy for each group of clusters. Due to the unbalanced dataset, the predictions could not be made for all clusters in the dataset. However, the analysis confirmed the observation that stop-initial clusters are produced significantly better than fricative-initial clusters. Additionally, the clusters composed of *stop* + *lateral* were more likely to be produced accurately than *stop* + *trill* clusters. Even though the sonority-based generalizations capture certain tendencies, in-depth analysis with the means of statistical models revealed that classifying segments by their manner of articulation, rather than by their sonority profiles, can make more detailed predictions regarding the order of acquisition of initial onset clusters.

The analysis also confirmed the observation that children who acquired more individual segments acquired more initial clusters. The results of the nested model comparison revealed that better predictions about accuracy can be made based on the degree of accuracy on singleton segments. In the next chapter, these results will be discussed with reference to theories of the phonological acquisition described in the theoretical background.

5. Discussion

5.1 Reduction patterns and the status of sC clusters

The results disclosed a discrepancy between simplification patterns targeting branching onsets and those targeting sC clusters. Branching onsets were produced with significantly greater accuracy than sC clusters. Therefore, fewer simplification processes targeted this group. This result is in accordance with the findings of previous studies on acquisition in Germanic languages that reported that branching onsets clusters are acquired before sC clusters (Fikkert & Freitas 2004, Yavaş 2008). Among branching onsets, the most common reduction pattern was the retention of an obstruent and the deletion of a more sonorous segment. Other studies also reported this observation (Pater 1997, Ohala 1999, Gnanadesikan 2004, Johnson & Reimers 2010, Barlow 2016). The only branching onset in which the more sonorous consonant was retained was the cluster /lj/. To account for these results, the mechanism underlying this pattern is usually attributed to the role of sonority, as is argued by Pater (1997), Ohala (1999), Gnanadesikan (2004) and Barlow (2016). Retaining the least sonorous consonant reshapes the syllable to have the greatest MSD from the onset consonant to the nucleus, which can be explained by a high-ranked markedness constraint against sonorants in an onset (Pater 1997, Gnanadesikan 2004).

Sonority is not the only possible driver of the reduction of a complex onset to an obstruent. This phenomenon could also be explained by the perception-based approach. As Vanderweide (2005) underlined, plosives have the strongest perceptual cues when they occur in the prevocalic position, due to the biggest difference in contrast between occurring segments. On the contrary, sonorants and vowels have more similar perceptual cues, making these two segments harder to distinguish. In addition, the perception-based account can explain why the /lj/ cluster was variably reduced to either /l/ or /j/.The variation could occur because both /l/ and /j/ are uttered as an approximant. Thus, these two phonemes have similar contextual cues, making them difficult to differentiate when they occur next to one another. Since these phonemes have comparable cues, children extracted /l/ in some cases, and /j/ in other cases.

Moreover, as reported by Másdóttir et al. (2021), sonorants that occur as the second member of a cluster, such as a trill or a lateral, are acquired later than plosive stops. Therefore, one could argue that children reduce clusters to stops because sonorants are mastered at a later stage. However, as Lleó & Prinz (1996) noted, cluster reduction does not always occur because an omitted segment has not been acquired yet, since a child may be able to produce it in another position in a word. This tendency was also observed in the data from Icelandic children, as shown by examples in (9). The fact that children produce the segment in another position in the same word could be due to perceptual cues –for instance, an approximant or a trill might be more challenging to perceive after a fricative, as opposed to an approximant or trill in the initial position of a cluster, or in the coda position.

(9). a. /**θv**ɔhtavjɛːl̥/ → [**θ**ɔttavjɛːl̥], ‘a washing machine’ (male, age 4;01)

b. /**θr**iːɹ̥/ → [**θ**iːɹ̥], ‘three’ (female, age 2;09)

In the current study, a cluster was categorized as produced accurately when a child produced two consonants in a cluster with segmental changes, rather than just one segment. Hence, the question arises as to why *stop* + *liquid* clusters were more frequently produced as two consonants with possible segmental changes, while *fricative* + *liquid* clusters were more frequently realized as one consonant. The possible explanation might also be attributed to perceptual cues: laterals and trills are harder to perceive after a fricative, and the cues of these phonemes are easier to extract after a plosive stop due to the greater difference in contrast.

When considering sC clusters, the simplification processes targeting this group varied between different classes of clusters. Overall, children produced the whole group of sC clusters less accurately than branching onsets. This finding indicates that sC clusters were more frequently produced as one segment compared to branching onsets. In branching onsets, reduction to the least sonorous consonant indicated the deletion of the second consonant (10a). Contrary to branching onsets, in *s* + *stop* and *s*+*nasal* clusters, the initial consonant was deleted. The group of *s*+*stop* clusters was mainly reduced to the stop, and the initial /s/ was deleted (10b). In the *s*+*nasal* clusters, this simplification pattern resulted in the retention of nasal, and deletion of the initial /s/ (10c).

(10). a. /**pl**ouːm/ → [**p**ouːm], ’a flower’ (male, age 2;05)

b. /**st**outl̥/ → [**t**outl̥], ‘a chair’ (male, age 3;06)

/**sp**ɪːl̥/ → [**p**ɪːl̥], ’a game’ (female, age 3;01)

c. /**sn**ʏːθ/ → [**n**ʏːθ], ‘a pacifier’ (male, age 2;07)

/**sm**ɛhkʏɹ̥/ → [**m**ɛhk], ‘a bib’ (female, age 3;05)

The children showed an equal degree of accuracy on *s* + *stop* clusters comprising sonority reversals, and *s* + *nasal* clusters comprising sonority rises. This finding indicates that reduction patterns in these clusters are not dependent on sonority. Regardless of the sonority class of the second cluster member, the initial /s/ was deleted. The same pattern was observed in triconsonantal sC clusters, where deletion of the initial /s/ was the most common simplification pattern (12a). In triconsonantal clusters, when the initial /s/ was deleted, the two retaining consonants were then reduced to the stop, which was the second-most common process that occurred in the whole group (12b).

(12). a. /**str**auːkʏɹ̥/ → [**tr**auːkʏɹ̥], ‘a boy’ (male, age 2;10)

/**stl**auŋka/ → [**tl**auŋka], ‘a snake’ (male, age 3;00)

b. /**str**auːkʏɹ̥/ → [**t**auːkʏθ], ‘a boy’ (female, age 3;11)

/**str**aiːtou/ → [**t**aiːtou], ‘a bus’ (male, age 2;08)

The deletion of the initial /s/ and retention of the sonorant (e.g., /str/ → /sr/), and retention of /s/ and the stop (e.g., /str/ → /st/) were rare processes targeting this group. Thus, children either deleted the initial /s/ and produced both following consonants (12a), or deleted both /s/ and the second consonant. This pattern might indicate that there are two stages in which children either produce the whole initial cluster and omit the initial /s/ before a stop, as in biconsonantal *s + stop* and *s + nasal* clusters, or produce only a stop followed by a vowel, as in the reduction of branching onsets.

Considering two other sC clusters, /sj/ and /sv/, there was a discrepancy in the level of accuracy – children produced /sv/ more frequently as one segment than /sj/. The /sv/ cluster matched other sC clusters in the accuracy level, while /sj/ seemed to pattern with branching onsets composed of an *obstruent + sonorant*. However, the investigation of simplification processes showed that the initial /s/ was retained or substituted with another segment (12a, 12b) and in the cluster /sv/, coalescence occurred as frequently as the retention of /s/ (12c).

(12). a. /**sj**ouːɹ̥/ → [**s**ouːɹ̥], ‘a sea*’* (female, age 2;06)

/**sv**iːn/ → [**s**iːn], ‘pork’ (female, age 2;07)

b. /**sj**ouːɹ̥/ → [**θj**ouːθ], ‘a sea’ (male, age 2;07)

/**sv**iːn/ → [**θv**iːn], ‘pork’ (female, age 3;07)

c. /**sv**iːn/ → [**f**iːn], ‘pork’ (male, age 3;00)

In these two clusters, the initial /s/ was often retained or substituted, contrary to *s+stop* and *s+nasal,* where the initial /s/ was deleted. Nevertheless, /sv/ was still produced with lower accuracy than /sj/. As discussed earlier, there is no consensus on whether the phonemes /v/ and /j/ should be analyzed as fricatives or approximants in Icelandic. From one perspective, it could be argued that /v/ should be analyzed as a fricative, and thus, the cluster /sv/ is composed of two consonants which share the manner of articulation. The results showed a discrepancy in the order of acquisition, wherein the cluster /sv/ was acquired later than the cluster /sj/. This discrepancy might indicate that only the phoneme /j/ should be analyzed as approximant. Clusters composed of two fricatives are acquired at the later stage, since producing them requires more motor control. Moreover, following Vanderweide’s proposal, the contrast between segments with the same manner of articulation is harder to perceive.

As noted above, the cluster /sv/ was frequently targeted by coalescence, which could have occurred due to the possibility of merging these two segments through their shared [+continuant] manner feature. Coalescence was also the most common simplification pattern that targeted the branching onset /θv/. The clusters /θv/ and /sv/ were most frequently produced as one segment /f/. The merged segment /f/ has the [labial] place feature from /v/, the laryngeal feature [-voice] from /θ/, and the manner feature [+continuant] from both segments. In these /sv/ and / θv/ clusters, the children’s incorporation of the voiceless feature of the initial /s/ and /θ/ into the preserved consonant might indicate that the initial fricative is more likely to be perceived in this position. The simplification patterns targeting these two clusters revealed that their reduction might not be driven by sonority. For these clusters, the children adopted the strategy of merging two segments instead of retaining the initial consonant.

The high occurrence of coalescence in the clusters with /v/ as a second consonant might reveal that Icelandic children analyze this phoneme as a fricative, whereas the phoneme /j/ patterns with other sonorants. Besides the discrepancy in the level of accuracy level in the production of /sv/ and /sj/, when /v/ occurred after a stop in the cluster /t^hv/, it was produced with a lower accuracy than the cluster /pj/ with /j/ as a second member. Hence, it can be argued that only the phoneme /j/ should be analyzed as approximants when they appear as the second member of a cluster, as in the phoneme classification proposed by Helgason (1991), Másdóttir & Strokes (2006), and Heimisdóttir (2015). However, in all these sources /v/ was also treated as an approximant.

Bjorndahl (2018) suggested that the phoneme /v/ might be analyzed as an obstruent when it appears as a first member of a cluster before a sonorant. The findings of the current study demonstrated that the clusters with /j/ as the second member (/mj/, /pj/, /sj/) were produced more frequently as two segments than the clusters with /v/ as the second member (/t^{h}v/, /sv/, /θv/). This observation. combined with the high occurrence of coalescence might indicate that the children treat the phoneme /v/ as a fricative, rather than an approximant. This is a topic that requires further investigation.

A similar discrepancy in reduction patterns of sC clusters was previously reported by Yavaş et al. (2008) and Krämer et al. (2017). In both studies, the children tended to delete /s/ when followed by a stop or a nasal, and retain /s/ when an approximant followed it. According to several proposals, *s+stop* clusters should be analyzed as appendix-initial or complex segments, while /s/ followed by a sonorant has the same tautosyllabic structure as branching onsets (Hall 1992, Weijer 1996). As Goad (2011) discussed, the simplification processes occurring in child language may provide additional evidence in support of this proposal. However, findings of the studies conducted by Yavaş et al. (2008), Krämer et al. (2017), and the study conducted in this thesis, revealed that the same reduction patterns that target *s+nasal* clusters also target *s+stop* clusters, , and do not target *s + sonorant* clusters. The same pattern was observed in the group of triconsonantal sC clusters – children either deleted the initial /s/ and reduced the two remaining consonants to the single stop, or they omitted the initial /s/ and produced both consonant. Based on these findings, it can be argued that /s/ is either an extrasyllabic element or more challenging for children to perceive in *s+nasal* and *s+stop* clusters. On the other hand, in clusters with the initial /s/ followed by an approximant, /s/ is either a part of a tautosyllabic onset or is easier to perceive in this position than before a nasal or a stop.

As Másdóttir et al. (2021) reported, the phoneme /s/ is acquired relatively late in Icelandic. However, when /s/ occurred before an approximant, it was more likely to be substituted than when it occurred before a nasal or a stop. Considering Vanderweide's (2005) proposal, it can be argued that the contextual cues of sibilant fricatives are easier to extract before an approximant. Thus, the children acquire *s + approximant* clusters before *s + nasal* and *s + stop.* It is beyond the scope of this study to conclude whether /s/ should be analyzed as an extrasyllabic element, and whether specific sC clusters have a non-tautosyllabic structure. Henke et al. (2012) argued that the SSP-violating *s+stop* clusters should not be treated as structurally different. According to them, treating sibilants as extrasyllabic does not explain why these phonemes, but not any other fricatives, are so frequently combined with plosives.

They argued that this type of consonant cluster is expected in the inventories of languages, since sibilants have stronger perceptual cues than any other fricatives before a plosive stop. Thus, following Vanderweide's (2005) proposal, sibilant fricatives followed by a stop should be acquired before other fricatives in this position. Icelandic does not have other *fricative + stop* initial clusters in its inventory, and therefore, this prediction could not be explicitly tested. Nonetheless, the current study demonstrated that children most frequently omitted the initial /s/ before a plosive and a nasal, and it was likely to be retained before an approximant. Thus, the findings of the current study suggested that /s/ is more likely to be perceived before an approximant, than before a stop or a nasal, which might be cause by the greater degree of contrast between cluster members in the latter. Therefore, to test the proposal of Henke et al. (2012), one should compare the acquisition order and the simplification patterns occurring in these two initial clusters in a language with different types of *fricative + stop* clusters.

To summarize, the investigation of simplification processes in child Icelandic revealed that different phonological processes target the groups of *obstruent + sonorant* and *s + approximant* clusters than *s + nasal* and *s + stop* clusters. In the former case, the cluster was usually reduced to the initial obstruent, while in the latter it was reduced to the second consonant, whether it was an obstruent or a sonorant. The observed discrepancy indicates that specific sC clusters diverge from other onset clusters structurally, or by perceptual cues. Additionally, the most common simplification processes that targeted branching onsets were coalescence and the retention of obstruents. Epenthesis and metathesis rarely occurred in data from Icelandic children. Garmann et al. (2020) suggested that the frequency in which reduction patterns occur can differ cross-linguistically, and reported that epenthesis was a frequent strategy in cluster simplification in Norwegian. The findings of the current study demonstrated that, in Icelandic, epenthesis is not a common reduction pattern, contrary to cluster reduction.

5.2 The effect of sonority-based markedness

A common assumption about acquisition in the OT literature is that children produce unmarked structures at the beginning stage of acquiring a grammar (Gnanadesikan 2004). As the course of the acquisition progresses, the constraint rankings shift, and children produce more complex structures. The second common assumption of OT is that there is continuity between child and adult grammar, and therefore, all patterns observed during acquisition will be found in languages (Pinker 1984). Given the continuity assumption, it has been argued that developmental stages in phonological acquisition reflect attested typologies of languages.

Hence, if markedness shapes the order of acquisition, its progression should be predicted based on the complexity level of segments and tautosyllabic clusters. The study conducted in this thesis investigated whether the order of acquisition of initial clusters can be predicted based on the implicational universals stated by the Sonority Dispersion Principle and the Minimal Sonority Distance.

According to the SDP, the most unmarked complex onsets are *obstruent + liquid* since these clusters have the most equally dispersed sonority from the first member of the onset to the nucleus (Clements 1990). On the other hand, the MSD favors clusters with the greatest sonority rise between the first and the second member of a cluster, i.e., *obstruents + glide/approximant* (Steriade 1982, Selkirk 1984). In the current study, I tested whether these two typologically motivated sonority-based generalizations about tautosyllabic onsets clusters are confirmed by the empirical data from child language. The MSD and the SDP are calculated based on the Sonority Hierarchy, which can be universal or language-specific. In general, assuming a different Sonority Hierarchy fundamentally changed the markedness status of clusters. Therefore, in the current study, both possibilities were investigated.

The results showed that when the MSD and SDP were calculated based on the fine-grained Sonority Hierarchy specific to Icelandic, none of the generalizations accounted for the order of acquiring tautosyllabic clusters. Also, the analysis showed that the SDP based on the more coarse-grained Sonority Hierarchy was not a significant predictor of accuracy. For the reasons discussed at the end of section 5.3., the predictions of these generalizations were not investigated further. Hence, only the predictions of the MSD based on the coarse-grained Sonority Hierarchy were thoroughly investigated.

The MSD states that clusters comprising sonority reversals are more marked than sonority plateaus. This assumption was not proven by the analyzed data since children produced sonority reversals with greater accuracy than sonority plateaus. As discussed in the previous section, the coalescence of clusters /sv/ and /θv/ was likely due to the possibility of merging these two phonemes, rather than their sonority profiles. Notably, the SD did not affect the accuracy in producing sC clusters: children produced *s+nasal* clusters with the SD of 1 with the same accuracy as the most marked *s+stop* clusters. The analysis confirmed that the group of sC clusters significantly differs from the group of branching onset, which indicated that Icelandic children acquire branching onsets before sC clusters. This order of acquisition was also reported for other Germanic languages by Yavaş et al. (2008).

Thus, the effect of MSD on branching onsets was investigated in isolation. The predictions of the SD about the most unmarked clusters were correct; children produced the clusters with the MSD of 3 and 4 the most accurately. Nevertheless, the MSD did not capture children's low accuracy on clusters with fricatives as the first member. As the analysis showed, the group of clusters with the SD of 1 were just as likely to be produced as accurately as the group with the MSD of 2, which included fricative-initial clusters Also, this group included the cluster /mj/, which children produced with significantly higher accuracy than fricative-initial clusters. This discrepancy could occur because nasals are one of the earliest acquired phonemes acquired in Icelandic. Even though the cluster /mj/ has a small rise in sonority, children produced this cluster with high level of accuracy, compared to the accuracy in *stop* + *lateral* clusters. Therefore, although the MSD explained some patterns observed in the data and the highest accuracy of *stop* + *liquid* and *stop* + *approximant*, this generalization did not capture children's low accuracy on *fricative* + *sonorant* clusters. According to the MSD, *fricative* + *liquid* clusters are less marked than *lateral* + *approximant* and *lateral* + *approximant* clusters. As shown in the comparison of the overall accuracy on clusters, these two types were produced with higher accuracy, and the statistical analysis found no significant difference between these two groups. The current study's results revealed that the markedness ranking of consonant clusters stated by the MSD and the SDP cannot capture the order of acquisition in the analyzed data. These generalizations did not capture the tendencies, which were predicted by the natural classes of individual segments of the clusters.

The results of the study conducted by Jarosz (2017) showed that Polish children were most accurate in the production of complex onsets with great sonority rises, and were least accurate on onsets with sonority reversals. Also, children's accuracy increased gradually with the rise in sonority. However, in Jarosz's (2017) study, sC clusters were not distinguished from branching onsets. In the supplementary materials to the study, frequencies of all investigated clusters are given. The most frequent sonority reversals in the analyzed dataset were *s* + *stop* clusters (/st/, /sp/, /sk/). The supplementary material also includes an analysis of children's error patterns, which aligns with the current study's findings. The results show that the most common simplification pattern for the class of *fricative* + *stop* was the deletion of the initial consonant, while when an obstruent was followed by a glide, or a nasal, the second consonant was deleted. The current study showed that investigating sC clusters together with branching onsets leads to different results. At first, the correlation between accuracy and the SD seemed to predict accuracy since children's accuracy increased with the SD of clusters. However, children's accuracy on sC clusters comprising sonority reversals was as low as their accuracy on sC clusters with sonority rises. The low accuracy in *s + nasal* clusters with the SD of 1 decreased the level of accuracy across the whole group. Nonetheless, when branching onsets were investigated separately, the accuracy on clusters with the MSD of 1 was higher than that in the group with the clusters with the SD of 2. Therefore, discrepancies in accuracy and simplification patterns targeting initial onsets lead to the conclusion that sC clusters and branching onsets should be investigated as two separate groups.

Furthermore, the findings showed that markedness of individual segments anticipated the order of acquisition of branching onsets with these segments. Scholars argue that the patterns observed in child phonology are an instance of the Emergence of the Unmarked (Fikkert 1994, Pater 1997, Goad & Rose 2004, Gnanadesikan 2004). As markedness optimizes fully developed grammar, it has been argued to be active at the beginning stage of acquisition. Therefore, it has been proposed that constraints that reinforce well-formed syllable shapes will also affect child grammar. Most of these proposals assume that children have adult-like mental representation, and the output of their grammar is derived from the segmentally accurate input. Smith (2010) argued that mispronunciation is caused by physiological limitations that affect both vocal tract control and comprehension. Therefore, Smith (2010) argued that patterns observed in child language are related to performance, rather than competence, the latter of which is understood as the mental representation of grammar. Whether patterns observed in child language are related to performance or competence, as well as the nature of children's mental representations are highly complex issues, which cannot be fully addressed in this thesis. I did, however, demonstrate that the patterns in child phonology can also be explained on the basis of markedness related to natural classes and their perceptual cues.

The Substance-Free approach to phonological theory rejects the notion of markedness, since it has been argued that child phonology is attributed to limitations in performance. In this theory, investigating child language and linking it to typological patterns is irrelevant (Reiss 2018). Nevertheless, the findings of studies discussed in this paper show that some typological patterns are linked to the acquisition order and the errors made by children, and that this association might not be trivial. The findings of Jarosz (2017) demonstrated that children's accuracy cannot be predicted based on the input frequency. Even though children frequently hear clusters composed of two obstruents, this does not indicate that these clusters will be acquired before typologically more frequent *obstruent + sonorant* clusters. Thus, this tendency could not be deduced from the language input. Everett & Schwartz (2023) demonstrated that typologically common labials are acquired before phonemes that occur in the ambient language more frequently. The findings of the current study revealed that clusters with fricatives are acquired later than clusters with more-typologically frequent stops. On the other hand, Vanderweide (2005), adhering to Licensing-by-Cues (Steriade 1999, 2009), argued that markedness is also linked to the perceptibility of segments in the onset clusters. Therefore, patterns observed in child language can be the result of articulatory, as well as perceptual markedness of segments and structures that occur in the ambient language.

Markedness relates to articulatory ease, which affects child language and shapes inventories of languages. The current study's findings indicated that markedness of individual segments, whether is related to articulatory or perceptual salience, can influence the order of acquisition of initial clusters. Considering this paper’s findings, it can be argued that markedness is linguistically relevant. It goes beyond the scope of this thesis to answer whether a constraint against the appendix is active in child grammar in the same way it is active in grammar of languages, or whether initial sibilants are simply hard to perceive, and are therefore deleted in child speech. Nevertheless, the analogy in the observed patterns might indicate that the same principles shape inventories of languages also affect phonological development. Therefore, markedness appears to be a relevant aspect of phonological theory.

5.3 The effect of type frequency

As many studies suggested, frequency plays a role in phonological development. Children who acquire different languages do not follow the same developmental path, and the order of acquisition of phonemes and syllable structure varies cross-linguistically. Even though frequency alone cannot account for all patterns in phonological development, it is expected to interact with other factors shaping phonological development (Edwards et al. 2015). This assumption was confirmed by studies conducted by Jarosz (2017) and Jarosz et al. (2017).

The current study aimed to compare the effect of frequencies with other factors that shape phonological development. Given the lack of child-directed speech corpora for Icelandic and transcribed corpora of spoken Icelandic, I tested the effect of the phonotactic probabilities of initial onset clusters. Nevertheless, the type frequencies of initial onsets, calculated based on the Icelandic lexicon, did not make correct predictions about the acquisition order. It may be that frequencies obtained from child-directed corpora might better reflect the input that children are exposed to. Jarosz et al. (2017) demonstrated that token frequencies, rather than type frequencies, make better predictions about the acquisition outcomes. On the other hand, Edwards (2010) reported that English children acquire a low-type frequency English phoneme /ð/ with a high token frequency at a later stage. However, the late acquisition of dental fricatives may be attributed to the articulatory difficulty of this phoneme. Additionally, /ð/ is a high-frequency stem from its presence in functional rather than lexical words, i.e., the definite determiners, pronouns, and complementizers. Edwards et al. (2015) proposed that frequency will interact with other factors during phonological development, which indicates that it should not be investigated in isolation. The findings of the current study did not find the effect of type-frequency to be significant, but more research here is warranted.

5.4 The order of acquiring singleton consonants and initial clusters

When children start to speak, they first produce syllables with the CV shape, before they acquire more complex syllables with consonant clusters. It is well-documented that children acquire singleton onsets before complex onsets, which was also reported for Icelandic by Másdóttir et al. (2021). The current study aimed to investigate which factors have the most significant impact on the acquisition course. Therefore, the indicator of children's accuracy on singleton consonants was investigated together with other factors influencing phonological acquisition.

The results revealed that accuracy on singleton consonants is the most impactful of all investigated factors. Thus, children who are further along the acquisition process, and have acquired more singletons, showed a greater accuracy on clusters. However, the statistical analysis showed that other factors also explain patterns observed in the data. Undoubtedly, children who mastered more individual segments produced more clusters. However, the analysis showed that other factors also influence the order of acquisition and can account for the patterns observed in the data from Icelandic children.

5.5 The effect of articulatory complexity of singleton consonants

The main finding of the current study suggested that the articulatory complexity of individual segments in clusters can account for the order of acquiring clusters containing these segments. This is evident in Icelandic children's acquisition of stop-initial clusters before fricative-initial clusters. Also, the findings indicated that clusters with laterals as a second member are acquired before clusters with trills. This finding was confirmed to be significant for stop-initial clusters, i.e., *stop + lateral* clusters were more likely to be produced accurately in relation to *stop + trill* clusters. For fricative-initial clusters, there was no significant difference between the group of *fricative + trill* and *fricative + lateral* clusters. However, *fricative + lateral* clusters were more likely to be produced as two segments as compared to *fricative + trill* clusters. Low accuracy on fricative-initial clusters can be attributed to the late acquisition of this natural class.

Additionally, Icelandic children acquire laterals before trills (Másdóttir 2008, Másdóttir et al. 2021). However, in stop-initial clusters, children were more likely to substitute /r/ with another approximant. In contrast, fricative-initial clusters were more likely to be produced as one segment. The perception-based approach can possibly explain this discrepancy. Due to the similarity of contextual and intensity cues, approximants are easier to perceive before a stop than before a fricative (Vanderweide 2005). Therefore, children might be more likely to omit /r/ and /l/ when it is preceded by a stop.

The current study's findings also suggested that the perceptual and motor limitations can explain phonological processes occurring in child language, and the order of acquisition, as stated by Stoel-Gammon & Sosa (2007) and Smith (2010). As discussed throughout this thesis, the patterns found in simplification processes, and the order of acquisition of consonant clusters were often previously attributed to sonority. However, onset clusters with a great rise in sonority usually include consonants that children acquire first: plosives, liquids, and glides. Sonority plateaus that are comprised of two consonants with the same manner of articulation require advanced motor-control skills. Additionally, the low accuracy on sonority plateaus can be attributed to the similarity of contextual cues in the sequence of two consonants with the same manner of articulation, which do not contrast in the stream of the airflow.

Moreover, the findings of the case study conducted by Kistanova (2021) reported that sonority-driven generalization did not capture high accuracy on sonority plateaus in the data from a Russian child that she examined. Children develop their set of skills at an individual pace, and thus, child-internal physical capacities will interact with linguistic factors during phonological development (Davis & Bedore 2013). In this thesis, I demonstrated that processes that occur in child language, which have previously been explained by the preference for sonority rises, can be also explained in terms of the articulatory complexity of cluster members, or the perceptibility of segments occurring in a cluster. The results demonstrated that markedness – either articulatory or/and perceptual – of individual segments explained more patterns in the data than sonority-based generalization.

5.6 Limitations and opportunities for future research

The main challenge for this study was an unbalanced dataset. Most of the analyzed observations included stop-initial and fricative-initial clusters followed by liquids. For this reason, I could not investigate the acquisition of other types of initial onsets, such as *nasal* + *approximant* and *lateral* + *approximant.* Moreover, based on the analyzed data I could not draw a conclusion about the order of acquiring *stop* + *approximant* clusters. Collecting data from children and transcribing them is a long, challenging, and complex process. Consequently, in a large number of studies investigating phonological acquisition, the data sample is usually restricted to a limited number of children. Testing theoretical proposals on a more extensive data sample decreases the likelihood of assuming incorrect premises. It was for these reasons that I decided to investigate data from the Másdóttir corpus, rather than conducting an independent study. Future research should strive to design studies aimed at directly testing theoretical proposals. It would be advantageous for these studies to use a larger data sample.

Furthermore, investigating spontaneous children's speech has the important advantage of examining more naturalistic acquisition data, as done in the studies by Jarosz (2017), and Jarosz et al. (2017). These two studies also demonstrated how different frequency measurements can lead to other predictions about acquisition outcomes. Thus, future research should also aim to explore this topic and by means of different methods of statistical modeling further investigate the effect of input frequency and its interaction with other factors in phonological acquisition. Additionally, including more detailed predictors of articulatory difficulty in the model, such as accuracy on individual consonants for each child, could offer more insight into the investigation of the articulatory difficulty effect.

This study demonstrated that the perception-based approach can potentially explain the simplification processes occurring in initial clusters, and the order of acquisition based on the perceptual cues of segments. However, the study did not explicitly test all its assumptions. Vanderweide (2005) established her proposal on data from 10 Dutch children, and therefore, the proposal focused only on accounting for the patterns observed in the investigated data. Future research should focus on elaborating measurable predictions of this approach, and test whether the theory of perceptual markedness is confirmed by simplification patterns and the order of acquisition. The effect of articulatory difficulty and perceptual markedness of segments should be investigated alongside the effect of sonority through statistical modeling. This could determine whether there is any explanatory power left for sonority effects after considering all these predictors.

I examine general tendencies in data from Icelandic children, rather than individual acquisition paths in this study. Future research should also focus on investigating longitudinal data, which can form the basis for testing different models of grammar. As discussed in this thesis, a plausible model of grammar must not only be able to account for possible developmental paths, but also to capture variation occurring within individuals. Elaboration of these models and testing of their predictions on empirical data requires further research.

6. Conclusion

Throughout this thesis, I discussed several factors that have been argued to influence the acquisition of consonant clusters and phonological processes occurring in children's grammar. The first aim of the current study was to investigate simplification patterns targeting initial clusters in the speech of children acquiring Icelandic. The finding revealed a difference between processes targeting two sC clusters: *s+stop* and *s+nasal*, and branching onsets. Children consistently reduced branching onsets to the obstruent, while in *s+stop* and *s+nasal* clusters the initial /s/ was deleted regardless of the sonority profile of the following consonant. This phenomenon has previously been reported by Yavaş et al. (2008) and Krämer et al. (2017). The group of *s + approximant* clusters did not follow the former pattern: in these clusters /s/ was either retained, substituted, or its voiceless feature was merged into the remaining consonant. The analysis confirmed that Icelandic children produced sC clusters with significantly lower accuracy than branching onsets. Thus, sC clusters were acquired later than branching onsets in Icelandic, and the discrepancy in the reduction patterns and children's accuracy indicated that these two groups require separate investigation and analysis. The finding also revealed that the phonemes /j/ and /v/ might have different statuses in Icelandic, as suggested by Árnason (2011) and Bjorndahl (2018).

The second aim of the current study was to investigate whether two sonority-based generalizations: the Minimal Sonority Distance, and the Sonority Dispersion Principle, can account for the order of the acquisition of initial onset clusters in Icelandic. In this study, the MSD and the SDP were calculated for two forms of the Sonority Hierarchy. The first was a fine-grained Sonority Hierarchy specific to Icelandic, proposed by Heimisdóttir (2015), and the second was the more course-grained general sonority scale. The markedness status of onset clusters, as determined by the MSD and SDP based on the Icelandic-specific Hierarchy, did not correlate with children's accuracy. The MSD based on the course-grained scale appeared to make better predictions. Therefore, the correlation between the MSD of initial clusters and children's accuracy was thoroughly investigated via quantitative analysis. The results of the analysis showed that separating sC clusters from branching onsets revealed that, although the MSD made correct predictions about the most and the least marked onsets, it did not capture children's low accuracy on fricative-initial clusters. Crucially, none of the sonority-based generalizations investigated in this study could account for this pattern. Fricative-initial clusters were most frequently realized as one consonant, despite comprising a rise in sonority and having equally dispersed sonority values. Nonetheless, the MSD and SDP predicted children's high accuracy on clusters with greater rises in sonority, such as *stop + liquid*. The findings demonstrated that this pattern could also be explained based on the articulatory complexity of individual segments in a cluster. The analysis confirmed that children produced stop-initial clusters more accurately than fricative-initial clusters. The analysis also revealed that stop + lateral clusters were produced with a significantly greater accuracy than *stop + trill* clusters. As Másdóttir (2008) and Másdóttir et al. (2021) reported, Icelandic children acquire stops and laterals earlier than trills and fricatives. These findings lead to the conclusion that the articulatory complexity, or markedness, of individual cluster members, can account for more tendencies observed in the data from Icelandic children. From another perspective, I demonstrated that the observed patterns could also be related to the perceptual cues of cluster members. Hence, the order of acquiring onset clusters can be related to the markedness based on the contrast between segments rather than their sonority profiles, as Vanderweide (2005) proposed. Additionally, the findings demonstrated that the accuracy in production of singleton segments was a robust predictor of accuracy in production of on initial clusters, and the effect of type frequency was not a significant predictor of children's accuracy.

To summarize, I showed that what has previously been explained by the role of sonority can be better accounted for by other factors playing a role in phonological acquisition, namely the acquisition order of individual segment classes, and their articulatory and perceptual difficulty. Further research on phonological acquisition should strive to resolve whether the order of acquisition, as well as strategies adopted by children to deal with complex structures, can be explained by articulatory and perceptual markedness of segments without reference to sonority. Finally, I believe that investigating phonological processes in child language contributes to understanding the language faculty and offers important insight into the phonology of an ambient language.

References

Ambridge, B., Kidd, E., Rowland, C., & Theakston, A. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42(2), 239-273.

Anttila, A. (1997). Deriving variation from grammar. In *Variation, change and phonological theory*. In Hinskens, F., Van Hout, R., and Wetzels, L. (eds). 35-68 (Vol. 146). John Benjamins Publishing Company.

Árnason, K. (2011). *The Phonology of Icelandic and Faroese* (The Phonology of the World's Languages). Oxford University Press.

Barlow J. A. (2001). A preliminary typology of initial clusters in acquisition. *Clinical linguistics & phonetics*, 15(1-2), 9–13.

Barlow, J. A. (2016). Sonority in acquisition: A review. In Ball, M. J. & Müller, N. (eds.), *Challenging sonority: Cross-linguistic evidence from normal and disordered language,* 295-336. Equinox.

Beckman, M. E., & Edwards, J. (2010). Generalizing over lexicons to predict consonant mastery. *Laboratory Phonology,* 1(2), 319-343.

Berent, I., Lennertz, T., & Rosselli, M. (2012). Universal phonological restrictions and language- specific repairs: Evidence from Spanish. *The Mental Lexicon,* 13, 275–305.

Bjorndahl, Ch. (2018). *A story of /v/: Voiced spirants in the obstruent-sonorant divide.* Doctoral Dissertation, Cornell University.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (ed.) *The handbook of phonological theory,* 206–244. Blackwell.

Boersma, P. & Hayes, B. (2001). Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry,* 32(1), 45–86. http://www.jstor.org/stable/4179137

Boersma, P. (1998). *Functional Phonology. Formalizing the Interaction between Articulation and Perceptual Drives*. PhD Dissertation, University of Amsterdam. The Hague, Holland: Academic Graphics. https://www.fon.hum.uva.nl/paul/diss/diss.html

Boersma, P., Levelt, C. (2003). Optimality theory and phonological acquisition. Annual Review of Language Acquisition, 3, 1–50. https://doi.org/10.1075/arla.3.03boe

Boysson -Bardies, B. de, Sagart, L., Hailé, P. & Durand, C. (1986). Acoustic investigation of cross-linguistic variability in babbling. Ln B. Lindblom & R. Zetterstrom (eds), *Precursors of early speech*. Stockton Press.

Bybee, J. (2010). A usage-based perspective on language. In *Language, Usage and Cognition* (1–13). Cambridge University Press. http://doi.org/10.1017/CBO9780511750526.001

Carlisle, R. (1998). The acquisition of onsets in a markedness relationship. A Longitudinal Study. *Studies in Second Language Acquisition,* 20(2), 245-260.

Carlisle, R. (2006). The sonority cycle and the acquisition of complex onsets. *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology*. Benjamins.

Carr, P. (1993). *Phonology* (Modern linguistics series). Macmillan Press.

Chabot, A. (2022). On substance and Substance-Free Phonology: Where we are at and where we are going. Canadian Journal of Linguistics/Revue Canadienne De Linguistique, 67(4), 429–443. http://doi.org/10.1017/cnj.2022.37

Chew, P. A. . (2003). *A computational phonology of Russian*. Dissertation.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. Praeger.

Clements, G. N. (1990). The role of the Sonority Cycle in Core Syllabification. Kingston & Beckman (eds). *Laboratory Phonology 1*. Cambridge University Press.

Clements, G. N. (2009). Does Sonority Have a Phonetic Basis? In E. Raimy, and Ch. E. Cairns (eds), *Contemporary Views on Architecture and Representations in Phonology*. MIT Press Scholarship. https://doi.org/10.7551/mitpress/9780262182706.003.0007

Clements, G. N., & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the Syllable*. MIT Press. http://dx.doi.org/10.1017/CBO9780511611834

Coetzee, A. W. (2016). A comprehensive model of phonological variation: grammatical and non-grammatical factors in variable nasal place assimilation. *Phonology*, 33(2), 211–246. http://doi.org/10.1017/S0952675716000117

Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. Phonology, 28(2), 197–234. http://www.jstor.org/stable/41303164

Davis, B. L., & Bedore, L. M. (2013). *An emergence approach to speech acquisition: Doing and knowing.* Taylor & Francis Group.

Demuth, K, (2011), The Acquisition of Phonology. In Goldsmith, J., Riggle, J., and Alan C. L. Yu (eds.) *The Handbook of Phonological Theory, 2nd ed*. Black-well. 571–595.

Demuth, K. (1995) Markedness and the development of prosodic structure. In J. Beckman (ed.), *Proceedings of the North Eastern Linguistic Society,* 25, 13-25. Amherst, MA: GLSA, University of Massachusetts.

Dodd, B., & Iacono, T. (1989). Phonological disorders in children: Changes in phonological process use during treatment. *British Journal of Disorders of Communication*, 24(3), 333–352. https://doi.org/10.3109/13682828909019894

Edwards, J., Beckman, M. E., & Munson, B. (2015). Frequency effects in phonological acquisition. *Journal of child language*, 42(2), 306–322. https://doi.org/10.1017/S0305000914000634

Everett, C. & Schwartz, S. (2023). The typological frequency of consonants is highly predictive of their order of acquisition in English. *Linguistic Typology*. https://doi.org/10.1515/lingty-2022-0033

Fikkert, P. & de Hoop, H. (2009). Language acquisition in optimality theory. *Linguistics* 47(2), 311-357. https://doi.org/10.1515/LING.2009.012

Fikkert, P. & Freitas, M. (2004). The role of language-specific phonotactics in the acquisition of onset clusters. *Linguistics in The Netherlands*. 21. 58-68. 10.1075/avt.21.09fik.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Vol. 6, pp. XIV, 358). Doctoral Dissertation. Holland Institute of Generative Linguistics. Rijksuniversiteit Leiden, Holland Academic Graphics.

Frisch, S. (2015). A preliminary investigation of quantitative patterns in sonority sequencing. *Rivista Di Linguistica, 27*(1), 9-27. https://www.italian-journal-linguistics.com/2015-2/

Garmann, N. G., Simonsen, H. G., Hansen, P., Holm, E., Post, B., & Payne, E. (2020). Cross-linguistic variation in word-initial cluster production in adult and child language: evidence from English and Norwegian. *Journal of Child Language*, 48(1), 1–30. http://doi.org/10.1017/S0305000920000069

Gerrits, E., & Zumach, A. (2006). The acquisition of# sC-clusters in Dutch. *Journal of Multilingual Communication Disorders*, 4(3), 218-230.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. J*ournal of Speech, Language, and Hearing Research*, 42(3), 708-26.

Gnanadesikan, A. (2004), Markedness and Faithfulness in Child Phonology. In R. Kager, J. Pater, W. Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge University Press.

Goad, H. (2011). The Representation of sC Clusters. In M. Oostendorp, C.J. Ewen, E. Hume & K. Rice (eds.). *The Blackwell Companion to Phonology*, 1-26. https://doi.org/10.1002/9781444335262.wbctp0038

Goad, H., & Rose, Y. (2004). Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. In Kager, R., Pater J., & Zonneveld, W. (eds.) Constraints in phonological acquisition, 109–157. Cambridge University Press. http://doi.org/10.1017/CBO9780511486418.005

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Blackwell.

Greenberg, J. (1978). Some generalizations concerning initial and final consonant clusters. In J. Greenberg, C. Ferguson, & E. Moravcsik (Eds.), *Universals of Human Language 2,* 243–279. Stanford University Press.

Gries, S. T. (2013). *Statistics for Linguistics with R*. De Gruyter, Inc.

Grunwell, P. (1987). *Clinical Phonology (2nd ed).* Croom Helm.

Hall, K. C., Blake A, Coates, E., Fry, M., Huang, S., Johnson, K., Lo, R., Scott M., Nam, S., McAuliffe, M. (2021). *Phonological CorpusTools*, Version 1.5.0. [Computer program]. Available from PCT GitHub page.

Hall, T. (1992). *Syllable structure and syllable-related processes in German* (Vol. 276, Linguistische Arbeiten). Niemeyer.

Hayes, B. P. (1999). Phonetically driven phonology. In Darnell, M., Moravcsik, E. A., Noonan, M., Newmeyer, F. J., & Wheatley, K. M. (Eds). *Functionalism and formalism in linguistics. Volume I: general papers.* 243-285. John Benjamins Publishing Company.

Heimisdóttir, L. Ö. 2014. Phonological opacity and Icelandic preaspiration. *University of Pennsylvania Working Papers in Linguistics 20(16).* https://repository.upenn.edu/pwpl/vol20/iss1/16/

Helgason , P. (1991). *On coarticulation and connected speech processes in Icelandic*. Unpublished script of MA thesis. University of Reading, UK .

Henke, E., Kaisse, E., & Wright, R. (2012). Is the Sonority Sequencing Principle an epiphenomenon? In *The Sonority Controversy* (Vol. 18), 65-100. De Gruyter. https://doi-org.mime.uit.no/10.1515/9783110261523.65

Hume, E. (2011). Markedness. In *The Blackwell Companion to Phonology* (eds M. Oostendorp, C.J. Ewen, E. Hume and K. Rice). John Wiley & Sons. https://doi.org/10.1002/9781444335262.wbctp0004

Jakobson, R. (1941/1968). *Child language, aphasia and phonological universals* (Reprint 2012. Ed., Vol. 72, Janua Linguarum. Series Minor). Mouton.

Jakobson, R. (1971). The sound laws of child language. In Jakobson, R. (ed.) *Studies on child language and aphasia,* 7– 20. Mouton. Reprinted in Waugh & Monville-Burston (1990), 294–304.

Jarosz, G. (2017). Defying the stimulus: acquisition of complex onsets in Polish. *Phonology,* 34(2), 269–298. http://doi.org/10.1017/S0952675717000148

Jarosz, G., Calamaro, Sh. & Zentz, J. (2017) Input frequency and the acquisition of syllable structure in Polish. *Language Acquisition*, 24:4, 361-399. Doi: 10.1080/10489223.2016.1179743

Jespersen, O. (1904). *Lehrbuch der Phonetik* [Textbook of phonetics]. Teubner.

Johnson, W., & Reimers, P. (2010). *Patterns in child phonology*. Edinburgh University Press.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants′ sensitivity to phonotactic patterns in the native language. *Journal of memory and Language*, 33(5), 630-645.

Kager, R. (1999). Learning OT grammars. In *Optimality Theory* (Cambridge Textbooks in Linguistics), 296-340. Cambridge University Press. doi:10.1017/CBO9780511812408.008

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Doctoral dissertation, MIT. http://hdl.handle.net/1721.1/16397

Kaye, J. (1992). Do you believe in magic? The story of s+ C sequences. *SOAS working papers in Linguistics and Phonetics*. Vol. 2, 293-313.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent Structure and Government in Phonology. *Phonology, 7(2),* 193–231. http://www.jstor.org/stable/4420016

Kent, Ray D. (1992). The biology of phonological development. In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds.), *Phonological development: Models, research, implications*, 65–90. York Press.

Kiparsky, P. (1979). Metrical Structure Assignment Is Cyclic. *Linguistic Inquiry, 10*(3), 421–441. http://www.jstor.org/stable/4178120

Kistanova, E. (2021). The Acquisition of Russian Word-Initial Consonant Clusters in a Russian-American English Bilingual Child: An Overview of One Child’s Exceptional Production. *Journal of Slavic Linguistics*, 29, 1. https://ojs.ung.si/index.php/JSL/article/view/164

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: results of functional data analysis. *The Journal of the Acoustical Society of America*, 124(5), 3158–3170. https://doi.org/10.1121/1.2981639

Krämer, M. (2018). Current issues and directions in Optimality Theory. In Bosch, A., & Hannahs, S. (eds.). *The Routledge Handbook of Phonological Theory* 1st ed., p. 37-67. Routledge.

Krämer, M., & Zec, D. (2020). Nasal consonants, sonority, and syllable phonotactics: The dual nasal hypothesis. *Phonology, 37*(1), 27-63. doi:10.1017/S0952675720000032

Krämer, M., Urek, O., & Markus, D. (2017). Word-initial consonant clusters in Latvian child language: variation, sonority, and grammar stratification. [Unpublished manuscript] https://www.semanticscholar.org/paper/Word-initial-consonant-clusters-in-Latvian-child-%3A-Kramer-Urek/c852d14d2141a51ab9b4959152222368736ffac9

Kristoffersen, K. E., & Simonsen, H. G. (2006). The acquisition of #sC clusters in Norwegian. *Journal of Multilingual Communication Disorders,* 4(3), 231-241.

Kuhl, P., Andruski, J., Chistovich, I., Chistovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages* (pp. XXI, 425). Blackwell.

Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. *Language acquisition,* 8(3), 237-264. https://doi.org/10.1207/S15327817LA0803_2

Levshina, N. (2015). *How to do Linguistics with R. Data exploration and statistical analysis*. John Benjamins Publishing Company.

Lleó, C., & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of child language*, 23(1), 31–56. https://doi.org/10.1017/s0305000900010084

Ludusan, B., Mazuka, R. & Dupoux, E. (2021). Does infant-directed speech help phonetic learning? A machine learning investigation. *Cognitive Science 45*. 1–31. https://doi.org/10.1111/cogs.12946.

Marecka, M., & Dziubalska-Kolaczyk, K. (2014). Evaluating models of phonotactic constraints on the basis of sC cluster acquisition data. *Language Sciences*, 46, 37-47. https://doi.org/10.1016/j.langsci.2014.06.002

Másdóttir, T. (2008). *Phonological development and disorders in Icelandic-speaking children* [Unpublished doctoral dissertation, University of Newcastle upon Tyne].

Másdóttir, T. (2014). *Málhljóðapróf ÞM [ÞM's Speech Sound Test]. A standardized phonological and articulation test for children at the age of 2;6-7;11 years of age*. Heyrnar og talmeinastöð Íslands [National Hearing and Speech Institute of Iceland]. doi: 10.13140/RG.2.1.4555.0486

Másdóttir, T., & Stokes, S. F. (2016). Influence of consonant frequency on Icelandic-speaking children's speech acquisition. *International journal of speech-language pathology,* 18(2), 111–121. https://doi.org/10.3109/17549507.2015.1060525

Másdóttir, T., McLeod, S., & Crowe, K. (2021). Icelandic Children’s Acquisition of Consonants and Consonant Clusters. *Journal of Speech, Language, and Hearing Research*, 64(5), 1490-1502.

McAllister Byun, T., & Tiede, M. (2017). Perception-production relations in later development of American English rhotics. *PloS one*, 12(2). https://doi.org/10.1371/journal.pone.0172022

McCarthy, J. & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistics Society 24*. 18. https://scholarworks.umass.edu/linguist_faculty_pubs/18

McCarthy, J.J. (2007), What Is Optimality Theory? *Language and Linguistics Compass*, 1: 260-291. https://doi.org/10.1111/j.1749-818X.2007.00018.x

McLeod, S., & Crowe, K. (2018). Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review. *American journal of speech-language pathology*, 27(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

McLeod, S., Doorn, J., & Reed, V. (2001). Normal Acquisition of Consonant Clusters. *American Journal of Speech-language Pathology*, 10(2), 99-110. https://www-proquest-com.mime.uit.no/intermediateredirectforezproxy

McLeod, S., Doorn, J., & Reed, V. A. (1997). Realizations of consonant clusters by children with phonological impairment. *Clinical linguistics & phonetics*, 11(2), 85–113. https://doi.org/10.1080/02699209708985185

Morelli, F. (1999). *The phonotactics and phonology of obstruent clusters in optimality theory*. Doctoral Dissertation, University of Maryland at College Park. https://doi.org/doi:10.7282/T37W6B13

Nikulásdóttir, A. B, Ármannsson, B., & Bergþórsdóttir, B. (2022). *Icelandic Pronunciation Dictionary for Language Technology 22.01*, CLARIN-IS, http://hdl.handle.net/20.500.12537/181.

Ohala, D. K. (1999). The influence of sonority on children's cluster reductions. J*ournal of Communication Disorders,* 32(6), 397–422. https://doi.org/10.1016/S0021-9924(99)00018-0

Parker, S. (2002). *Quantifying the Sonority Hierarchy*. Doctoral Dissertation. University of Massachusetts Amherst. https://scholarworks.umass.edu/dissertations/AAI3056268

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics.* 36. 55-90. https://doi.org/10.1016/j.wocn.2007.09.003

Parker, S. (2011). Sonority. In M. Oostendorp, C.J. Ewen, E. Hume and K. Rice (Eds). *Companions to Linguistics: The Blackwell Companion to Phonology, 2*, 1184. https://doi.org/10.1002/9781444335262.wbctp0049

Parker, S. (2012). Sonority distance vs. sonority dispersion – a typological survey. In S. Parker (Ed.), *The Sonority Controversy*, 101-166. De Gruyter Mouton. https://doi.org/10.1515/9783110261523.101

Pater, J. (1997). Minimal Violation and Phonological Development. *Language Acquisition*. 6:3, 201-253.

Pater, J., & Barlow, J. A. (2003). Constraint conflict in cluster reduction. *Journal of Child Language*, *30*(3), 487–526. https://doi.org/10.1017/S0305000903005658

Pinker, S. (1984). *Language learnability and language learning*. Cambridge, MA: Harvard University Press.

Prince, A. S. & Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar.* Rutgers and Boulder: University of Rutgers and University of Colorado, MS. Blackwell.

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reiss, C., & Volenec, V. (2022). Conquer primal fear: Phonological features are innate and substance-free. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique,* 67(4), 581–610. http://doi.org/10.1017/cnj.2022.35

Reiss, C., (2018). Substance Free Phonology. In Bosch, A., & Hannahs, S., (Eds.) *The Routledge Handbook of Phonological Theory* (First ed., Routledge handbooks in linguistics). Taylor and Francis.

Rice, K. (2007). Markedness in phonology. In P. Lacy (Ed.), *The Cambridge Handbook of Phonology* (Cambridge Handbooks in Language and Linguistics, 79-98). Cambridge University Press. Doi:10.1017/CBO9780511486371.005

Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain and language*, 64(1), 83–121. https://doi.org/10.1006/brln.1998.1958

Romani, C., Galuzzi, C., Guariglia, C., & Goslin, J. (2017). Comparing phoneme frequency, age of acquisition, and loss in aphasia: Implications for phonological universals. *Cognitive Neuropsychology*, 34:7-8, 449-471. doi: 10.1080/02643294.2017.1369942

Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In Durand J., Gut

U., & Kristoffersen, G. (Eds.), *The Oxford handbook of corpus phonology* (pp. 380-401). Oxford University Press.

Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In Durand, J., Gut, U., Kristoffersen, G. (eds). *The Oxford Handbook of Corpus Phonology (1st ed.).* Oxford University Press.

Saussure, F.de, (1916). *Cours de linguistique Générale* [General linguistics course]. Payot.

Scheer, T. (2022). 3 x Phonology. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique*, 67(4), 444–499. http://doi.org/10.1017/cnj.2022.22

Selkirk, E. (1982). The Syllable. In H. Hulst & N. Smith (Ed.), *The Structure of Phonological Representations, Part 2* (337-384). De Gruyter. https://doi.org/10.1515/9783112423325-010

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Mass, MIT press.

Shatz, I. (2019). Phonological selectivity in the acquisition of English clusters. *Journal of Child Language*, 46(6), 1025-1057. Doi:10.1017/S0305000919000345

Shriberg, L., Austin, D., Lewis, B., McSweeny, J., & Wilson, D. (1997). The Percentage of Consonants Correct (PCC) Metric: Extensions and Reliability Data. *Journal of Speech, Language, and Hearing Research*, 40(4), 708-722. https://doi.org/10.1044/jslhr.4004.708

Sievers, E. (1881). *Grundzüge der Phonetik* [Fundamentals of phonetics]. Breitkopf & Härtel.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *The Journal of speech and hearing disorders*, 55(4), 779–798. https://doi.org/10.1044/jshd.5504.779

Smit, A.B. (1993). Phonological errors distributions in the Iowa-Nebraska articulation norms project: word-initial consonant clusters. *Journal of Speech and Hearing Research*. 36: 931–947. https://doi.org/10.1044/jshr.3605.931

Smith, N. (2010). *Acquiring phonology: A cross-generational case-study* (Vol. 124, Cambridge Studies in linguistics). Cambridge University Press.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. Doctoral dissertation, Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/15653

Steriade, D. (1999). Alternatives to the syllabic interpretation of consonantal phonotactics. In O. Fujimura, B. Joseph and B. Palek (eds), *Proceedings of the 1998 Linguistics and Phonetics Conference*. The Karolinum Press, 205-242.

Steriade, D. (2001). Directional Asymmetries in Place Assimilation. In B. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology,* Academic Press, 219-250.

Steriade, D. (2009). The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In: Hanson, K. and Inkelas, S. (eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 151–179. Cambridge, MA: MIT Pres https://doi.org/10.7551/mitpress/9780262083799.003.0007

Stites, J., Demuth, K. & Kirk. C. (2004). Markedness vs. frequency effects in coda acquisition. In Brugos, A., Micciulla, L. & Smith. Ch. E. (eds.) *Proceedings of the 28th Annual Boston University Conference on Language Development*, 565–576. Cascadilla Press.

Stoel-Gammon, C. & Sosa, A. V. (2007). Phonological development. In E. Hoff & M. Schatz (eds), *Handbook of child language*, 238–56. Blackwell Publishing Ltd. https://doi.org/10.1002/9780470757833.ch12

Stoel-Gammon, C. (1987). Phonological skills of 2-year-olds. *Language, Speech, and Hearing Services in Schools*, 18(4), 323–329. https://doi.org/10.1044/0161-1461.1804.323

Storkel H. L. (2018). The Complexity Approach to Phonological Treatment: How to Select Treatment Targets. *Language, speech, and hearing services in schools*, 49(3), 463–481. https://doi.org/10.1044/2017_LSHSS-17-0082

Tesar, B., & Smolensky, P. (1998). Learnability in Optimality Theory. *Linguistic Inquiry*, 29(2), 229–268. http://www.jstor.org/stable/4179017

Thráinsson, H., & Gíslason, I. (1993). *Handbók um íslenskan framburð* [Handbook on Icelandic pronunciation]. Rannsóknarstofnun Kennaraháskóla Íslands.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29(1), 66–85.

Trubetzkoy, N. S., (1939/1971). *Grundzüge der Phonologie* (5. Aufl. ed.) [Fundamentals of phonology]. Vandenhoeck & Ruprecht.

Van de Weijer, J. (2017). Emergent Phonological Constraints: The acquisition of *COMPLEX in English. *Acta Linguistica Academica*, 64(1), 153–165. https://www.jstor.org/stable/26452346

Vanderweide, T. (2005). The acquisition of manner in pre-vocalic sequences: a cue is a cue. In M. Tzakosta, C.C. Levelt and J.M. van de Weijer (Eds), *Developmental Paths in Phonological Acquisition*. Special issue of Leiden Papers in Linguistics 2.1 (2005), 137-161. http://www.ulcl.leideniniv.nl

Vaux, B., & Wolfe, A. (2009). The Appendix. In Raimy, E. & Cairns, Ch. E. (eds), *Contemporary Views on Architecture and Representations in Phonology*. The MIT Press Scholarship Online. https://doi.org/10.7551/mitpress/9780262182706.001.0001,

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin* (Reprint 2011. Ed.). Mouton de Gruyter.

Vitevitch, M.S. & Luce, P.A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers*, 36, 481-487. https://link.springer.com/article/10.3758/BF03195594

Weijer, J. (1996). *Segmental structure and complex segments* (Vol. 350, Linguistische Arbeiten). Max Niemeyer Verlag.

Winter, B. (2019). *Statistics for linguists : an introduction using R*. Routledge.

Yavaş, M. (2006). Sonority and the acquisition of #sC clusters, *Journal of Multilingual Communication Disorders*, 4:3, 159-168. doi: 10.1080/14769670601110473

Yavaş, M. (2013). What explains the reductions in/s/-clusters: Sonority or [continuant]?. *Clinical linguistics & phonetics*, 27(6-7), 394-403. https://doi.org/10.3109/02699206.2013.767378

Yavaş, M. (2014). What guides children's acquisition of #sC clusters? *Perspectives on phonological theory and development: In honor of Daniel A. Dinnsen*, 56, 115. https://doi.org/10.1075/lald.56.11yav

Yavaş, M., Avivit, B.D., Gerrits, E., Kristoffersen, K. E., & Simonsen, H., G. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics* 22(6), 421-441. doi: 10.1080/02699200701875864

Zamuner, T. S., Gerken, L. & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language* 31(3). 515–536. https://doi.org/10.1017/S0305000904006233

Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology* (161–194). Cambridge University Press.

Zhao, X., & Berent, I. (2016). Universal Restrictions on Syllable Structure: Evidence from Mandarin Chinese. *Journal of psycholinguistic research*, *45*(4), 795–811. https://doi.org/10.1007/s10936-015-9375-1

Appendix I - The dataset.

Words with initial consonant clusters from the from the Malhjodaprof, available as a part of Másdóttir (2014) corpus, included in the dataset.

Table (1). Biconsonantal clusters

	Orthographic form	**Target transcription**	**Gloss**	**Initial cluster**
1	blóm	[ˈplouːm]	*a flower*	/pl/
2	ljón	[ˈljouːn]	*a lion*	/lj/
3	bleia	[ˈpleija]	*a diaper*	/pl/
4	gluggi	[ˈklʏccɪ]	*a window*	/kl/
5	skip	[ˈscɪːp]	*a boat*	/sc/
6	greiða	[ˈkreiːða]	*a comb*	/kr/
7	spil[14]	[ˈspɪːl̥]	*a game*	/sp/
8	brauð	[ˈprœyːθ]	*a bread*	/pr/
9	fluga	[ˈflʏːɣʏɹ̥]	*a fly*	/fl/
10	skæri	[ˈscaiːrɪ]	*scissors*	/sc/
11	prinsessa	[ˈpʰrɪnsɛssa]	*a princess*	/pʰr/
12	plástur	[ˈpʰlaustʏɹ̥]	*a band-aid*	/pʰl/
13	blása	[ˈplauːsa]	*to blow/ blowing*	/pl/
14	brú	[ˈpruː]	*a bridge*	/pr/
15	tromma	[ˈtʰrɔmma]	*a drum*	/tʰr/
16	tveir	[ˈtʰveiːɹ̥]	*two*	/tʰv/
17	draugur	[ˈtrœyːɣʏɹ̥]	*a ghost*	/tr/
18	krummi	[ˈkʰrʏmmɪ]	*a raven*	/kʰr/
19	klukka	[ˈkʰlʏhka]	*a clock*	/kʰl/
20	grís	[ˈkriːs]	*a pig*	/kr/

---

[14] The word [spɪːl̥] was included twice in the test.

	Orthographic form	**Target transcription**	**Gloss**	**Initial cluster**
22	froskur	[ˈfrɔskʏɹ̥]	*a frog*	/fr/
23	flauta	[ˈflœyːta]	*a flute*	/fl/
24	spil[3]	[ˈspɪːl̥]	*a game*	/sp/
25	stóll	[ˈstoutl̥]	*a chair*	/st/
26	skór	[ˈskouːɹ̥]	*a shoe*	/sk/
27	skegg	[ˈscɛkː]	*a beard*	/sc/
28	smekkur	[ˈsmɛhkʏɹ̥]	*a bib*	/sm/
29	snuð	[ˈsnʏːθ]	*a pacifier*	/sn/
30	sjór	[ˈsjouːɹ̥]	*a sea*	/sj/
31	svín	[ˈsviːn]	*pork*	/sv/
32	þrír	[ˈθriːɹ̥]	*three*	/θr/
33	bjalla	[ˈpjatla]	*a bell*	/pj/
34	þvottavél	[ˈθvɔhtavjɛːl̥]	*a washing machine*	/θv/
35	stelpa	[ˈstɛl̥pa]	*a girl*	/st/
36	mjólka	[ˈmjoul̥ka]	*to milk/milking*	/mj/

Table (2). Triconsonantal clusters:

	Orthographic form	**Target transcription**	**Gloss**	**Initial cluster**
1	slanga	[ˈstlauŋka]	*a snake*	/stl/
2	skríða	[ˈskríða]	*to crawl/ crawling*	/skr/
3	strætó	[ˈstraiːtou]	*a bus*	/str/
4	strákur	[ˈstrauːkʏɹ̥]	*a boy*	/str/

Table (3). Number of raw observations from each subject.

	Child ID	**Number of observations**		**Age (years;months)**	**Sex**	**Accuracy Score**
		CCV	**CCCV**			
1.	01-21	36	4	3;02	M	54%
2.	01-23	36	4	2;08	F	85%
3.	01-24	36	4	3;03	F	86%
4.	01-26	36	4	3;01	F	52%
5.	01-27	34	4	2;07	F	82%
6.	01-29	34	4	2;07	M	68%
7.	01-30	35	4	2;06	M	73%
8.	01-31	35	4	2;07	M	90%
9.	01-33	36	4	2;09	F	77%
10.	03-31	36	4	3;04	F	82%
11.	03-33	24	4	2;07	M	52%
12.	03-34	35	4	4;08	M	92%
13.	03-36	36	4	2;11	F	81%
14.	03-37	36	4	3;07	M	80%
15.	03-38	36	4	3;03	F	77%
16.	03-39	36	4	3;05	F	86%
17.	04-18	36	4	3;06	F	64%
18.	05-03	36	4	2;07	M	74%
19.	05-18	36	4	3;02	M	66%
20.	05-19	36	4	3;05	F	75%
21.	05-20	36	4	2;09	F	73%
22.	05-27	36	4	3;08	M	73%
23.	05-29	36	4	3;06	M	64%
24.	05-46	33	4	2;09	F	58%
25.	06-01	33	4	2;06	M	64%
26.	06-08	36	4	2;08	M	61%
27.	06-11	36	4	4;02	M	91%
28.	06-13	35	4	4;01	F	76%
29.	06-33	35	4	2;09	M	90%
30.	06-37	36	4	2;07	M	97%
31.	06-38	36	4	2;07	F	83%
32.	06-41	36	4	3;00	M	72%
33.	07-18	35	4	2;10	M	73%
34.	07-19	36	4	2;09	F	58%
35.	07-25	34	4	2;07	M	86%
36.	08-10	35	4	3;09	F	65%
37.	08-11	36	4	3;10	F	76%
38.	09-01	36	4	2;10	M	73%
39.	09-02	35	4	2;07	M	46%
40.	09-03	36	4	2;9	M	66%
41.	09-04	36	4	2;9	M	56%

42.	09-05	36	4	2;10	M	93%
43.	09-06	36	4	2;06	F	86%
44.	10-04	36	4	3;02	M	61%
45.	10-05	36	4	2;11	F	67%
46.	12-02	36	4	3;03	F	82%
47.	13-02	36	4	3;04	M	92%
48.	13-03	36	4	3;05	M	96%
49.	13-05	36	4	2;11	M	66%
50.	13-06	35	4	2;08	F	78%
51.	13-09	36	4	2;10	F	94%
52.	13-10	36	4	2;08	F	59%
53.	13-12	36	4	3;06	M	69%
54.	15-02	36	4	3;02	M	77%
55.	15-03	35	4	3;01	F	44%
56.	15-04	36	4	3;00	M	61%
57.	15-05	33	4	2;10	M	51%
58.	15-06	33	2	3;00	F	62%
59.	15-08	36	4	4;01	M	86%
60.	15-09	36	4	3;01	F	92%
61.	16-06	36	4	3;03	M	87%
62.	16-08	36	4	3;01	M	85%
63.	16-09	36	4	3;02	M	75%
64.	16-10	36	4	3;00	M	54%
65.	17-02	36	4	3;01	M	76%
66.	17-15	36	4	3;01	M	67%
67.	26-02	36	4	4;00	M	72%
68.	26-04	36	4	4;03	M	64%
Total		**2408**	**267**	**= 2675**		

Table (4). Raw and relative frequencies of branching onsets in the dataset.

Cluster	Raw frequency	Relative frequency in the dataset
/fl/	*133*	*5,5%*
/fr/	*68*	*2,8%*
/kʰl/	*68*	*2,8%*
/kʰr/	*67*	*2,8%*
/kl/	*132*	*5,5%*

/kr/	*132*	*5,5%*
/lj/	*67*	*2,8%*
/mj/	*65*	*2,7%*
/pʰl/	*68*	*2,8%*
/pʰr/	*67*	*2,8%*
/pj/	*67*	*2,8%*
/pl/	*202*	*8,4%*
/pr/	*134*	*5,6%*
/tʰr/	*68*	*2,8%*
/tʰv/	*67*	*2,8%*
/tr/	*67*	*2,8%*
/θr/	*66*	*2,7%*
/θv/	*67*	*2,8%*

Table (5). Raw and relative frequencies of sC clusters in the dataset.

Cluster	Raw frequency	Relative frequency in the dataset
/sj/	*67*	*2,8%*
/sv/	*65*	*2,7%*
/sc/	*200*	*8,3%*
/sk/	*68*	*2,8%*
/sp/	*136*	*5,6%*
/st/	*136*	*5,6%*
/sm/	*66*	*2,7%*
/sn/	*65*	*2,7%*

Table (6). Raw frequencies of sCC clusters (not included in the analyzed dataset)5.

Cluster	Raw frequency
/stl/	*134*
/str/	*66*
/skr/	*67*

Appendix II - Output of the regression analysis

1. The model with the MDS and SDP as numerical variables

(fitted to the dataset without sonority reversals and plateaus for which sonority dispersion cannot account)

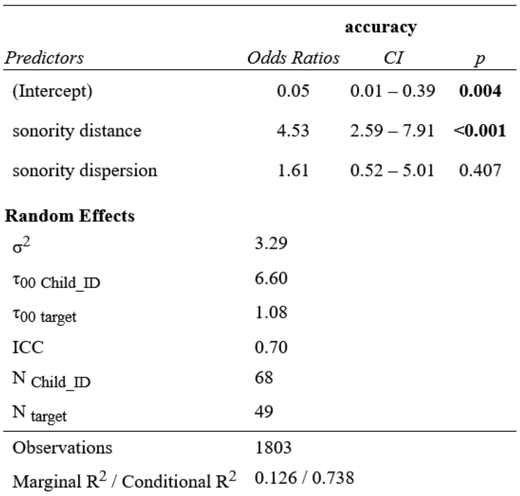
```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial  ( logit )
Formula: accuracy ~ 1 + sonority_distance + sonority_dispersion + (1 |  Child_ID) + (1 | target)
Data: clusteranalysis

     AIC      BIC   logLik deviance df.resid
  1470.3   1497.8   -730.2   1460.3     1798

Scaled residuals:
    Min      1Q  Median      3Q     Max
-6.4082 -0.3206  0.1431  0.3787  5.2176

Random effects:
 Groups   Name        Variance Std.Dev.
 Child_ID (Intercept) 6.603    2.570
 target   (Intercept) 1.085    1.042
Number of obs: 1474, groups:  Child_ID, 68; target, 36


Fixed effects:
                    Estimate Std. Error  z value Pr(>|z|)
(Intercept)          -3.0204     1.0580   -2.855  0.00431 **
sonority_distance     1.5108     0.2843    5.313 1.08e-07 ***
sonority_dispersion   0.4787     0.5778    0.828  0.40743
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Test for overdispersion         chisq      ratio          rdf             p
                          1192.9818624  0.6635049 1798.0000000    1.0000000
```

```
Vif score:    sonority_distance    sonority_dispersion
                      1.691209              1.691209
```

C.score: 0.93

Log odds converted into the ratio of odds:



Predictors	Odds Ratios	CI	p
	accuracy		
(Intercept)	0.05	0.01 – 0.39	**0.004**
sonority distance	4.53	2.59 – 7.91	**<0.001**
sonority dispersion	1.61	0.52 – 5.01	0.407
Random Effects			
σ^2	3.29		
$\tau_{00\ \mathrm{Child\_ID}}$	6.60		
$\tau_{00\ \mathrm{target}}$	1.08		
ICC	0.70		
$N_{\mathrm{Child\_ID}}$	68		
N_{target}	49		
Observations	1803		
Marginal R^2 / Conditional R^2	0.126 / 0.738		

2. The model with sC clusters and branching onsets

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial  ( logit )
Formula: accuracy ~ 1 + segment_accuracy + sonority_distance + (1 | Child_ID) +      (1 | target)
   Data: clusteranalysis
     AIC      BIC   logLik deviance df.resid
  1904.9   1957.0   -943.5   1886.9     2402
Scaled residuals:
    Min      1Q  Median      3Q     Max
-7.6772 -0.3720  0.0662  0.3593 19.6153
Random effects:
 Groups   Name        Variance Std.Dev.
 Child_ID (Intercept) 6.8140   2.6104
 target   (Intercept) 0.6201   0.7875
Number of obs: 2411, groups:  Child_ID, 68; target, 64
```

Fixed effects (sonority distance -1 is the reference level):[15]

```
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)         -1.54381    0.44410  -3.476 0.000509 ***
segment_accuracy     0.03558    0.01392   2.556 0.010587 *
sonority_distance0  -1.57112    0.71611  -2.194 0.028238 *
sonority_distance1   0.65183    0.48820   1.335 0.181824
sonority_distance2   1.16266    0.49423   2.352 0.018650 *
sonority_distance3   3.55706    0.38406   9.262  < 2e-16 ***
sonority_distance4   3.01754    0.92085   3.277 0.001049 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

VIF score:

```
                      VIF  Df    VIF^(1/(2*Df))
segment_accuracy  1.15009  1       1.072422
sonority_distance 1.15009  5       1.014082
```

Test for overdispersion:

```
        chisq        ratio          rdf          p
 2422.7698901    1.0086469 2402.0000000    0.378883
```

C.score: 0.94

Predicted probabilities and the pairwise comparison of all levels of *sonority distance*:

```
$`emmeans of sonority_distance`
 sonority_distance   prob     SE  df asymp.LCL asymp.UCL
 -1                0.1904 0.0686 Inf    0.0895     0.360
 0                 0.0466 0.0322 Inf    0.0117     0.168
 1                 0.3110 0.1071 Inf    0.1449     0.546
 2                 0.4293 0.1208 Inf    0.2225     0.664
 3                 0.8919 0.0373 Inf    0.7945     0.946
 4                 0.8278 0.1316 Inf    0.4405     0.967
Confidence level used: 0.95, Intervals are back-transformed from the logit scale
```

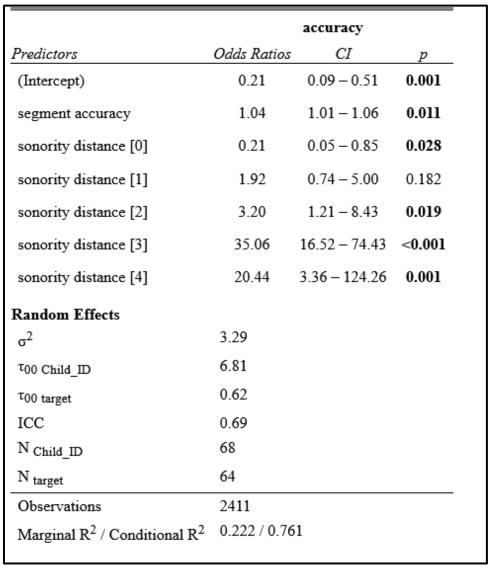
---

[15] Estimates are given in log odds.

The pairwise differences of sonority_distance (not significant pairings are bolded):

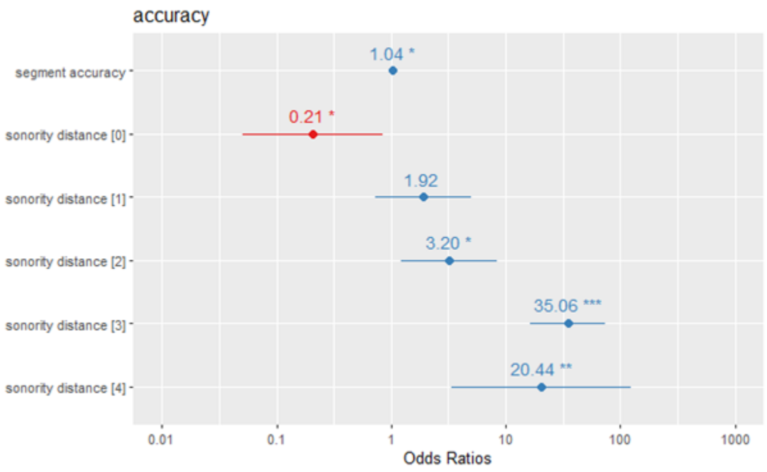
```
                                              estimate  SE   df   z.ratio p.value

 (sonority_distance-1) - sonority_distance0   1.571 0.716 Inf   2.194  0.0282
 (sonority_distance-1) - sonority_distance1  -0.652 0.488 Inf  -1.335  1.0000
 (sonority_distance-1) - sonority_distance2  -1.163 0.494 Inf  -2.352  0.0186
 (sonority_distance-1) - sonority_distance3  -3.557 0.384 Inf  -9.262  <.0001
 (sonority_distance-1) - sonority_distance4  -3.018 0.921 Inf  -3.277  0.0157
 sonority_distance0 - sonority_distance1     -2.223 0.752 Inf  -2.957  0.0465
 sonority_distance0 - sonority_distance2     -2.734 0.758 Inf  -3.604  0.0047
 sonority_distance0 - sonority_distance3     -5.128 0.694 Inf  -7.384  <.0001
 sonority_distance0 - sonority_distance4     -4.589 1.087 Inf  -4.221  0.0004
 sonority_distance1 - sonority_distance2     -0.511 0.544 Inf  -0.938  1.0000
 sonority_distance1 - sonority_distance3     -2.905 0.445 Inf  -6.527  <.0001
 sonority_distance1 - sonority_distance4     -2.366 0.947 Inf  -2.497  0.1878
 sonority_distance2 - sonority_distance3     -2.394 0.442 Inf  -5.412  <.0001
 sonority_distance2 - sonority_distance4     -1.855 0.948 Inf  -1.958  0.7543
 sonority_distance3 - sonority_distance4      0.540 0.890 Inf   0.606  1.0000
```

Log odds converted into the ratio of odds (the reference level is sonority distance -1):



	accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.21	0.09 – 0.51	**0.001**
segment accuracy	1.04	1.01 – 1.06	**0.011**
sonority distance [0]	0.21	0.05 – 0.85	**0.028**
sonority distance [1]	1.92	0.74 – 5.00	0.182
sonority distance [2]	3.20	1.21 – 8.43	**0.019**
sonority distance [3]	35.06	16.52 – 74.43	**<0.001**
sonority distance [4]	20.44	3.36 – 124.26	**0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{Child\_ID}}$	6.81		
$\tau_{00\ \text{target}}$	0.62		
ICC	0.69		
$N_{\text{Child\_ID}}$	68		
N_{target}	64		
Observations	2411		
Marginal R^2 / Conditional R^2	0.222 / 0.761		

A plot of main effects:

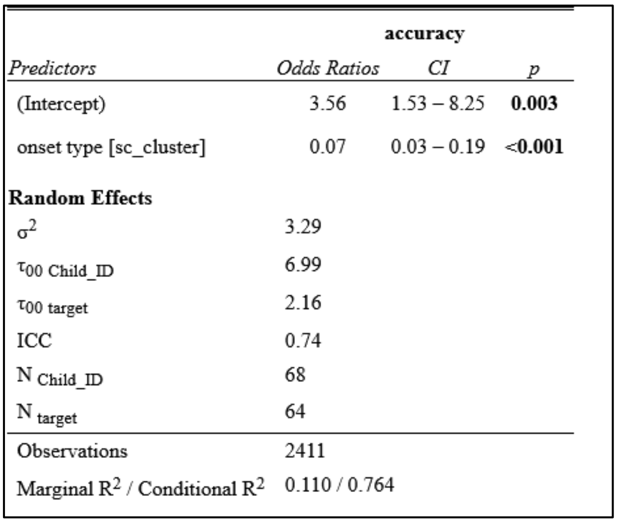


3. The model comparing branching onsets and sC clusters

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: accuracy ~ 1 + onset_type + (1 | Child_ID) + (1 | target)
   Data: clusteranalysis
    AIC      BIC   logLik deviance df.resid
  1942.0   1965.1   -967.0   1934.0     2407
Scaled residuals:
    Min      1Q  Median      3Q     Max
-7.9043 -0.3705  0.0687  0.3697 20.2009
Random effects:
 Groups   Name        Variance Std.Dev.
 Child_ID (Intercept) 6.986    2.643
 target   (Intercept)    2.157    1.469
Number of obs: 2411, groups:  Child_ID, 68; target, 64
Fixed effects (branching onset is the reference level):
                          Estimate Std. Error z value  Pr(>|z|)
(Intercept)                  1.2692     0.4290   2.959  0.00309 **
onset_typesc_cluster  -2.6232     0.4804  -5.460 < 4.75e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

C score: 0.94

Log odds converted into the ratio of odds:



	accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	3.56	1.53 – 8.25	**0.003**
onset type [sc_cluster]	0.07	0.03 – 0.19	**<0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{Child\_ID}}$	6.99		
$\tau_{00\ \text{target}}$	2.16		
ICC	0.74		
$N_{\text{Child\_ID}}$	68		
N_{target}	64		
Observations	2411		
Marginal R^2 / Conditional R^2	0.110 / 0.764		

4. The model including only branching onsets

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: accuracy ~ sonority_distance + segment_accuracy + (1 | Child_ID) +      (1 | target)
   Data: branchingonsets

     AIC      BIC   logLik deviance df.resid
  1209.2   1252.3   -596.6   1193.2     1601

Scaled residuals:
    Min      1Q  Median      3Q     Max
-9.1381 -0.2213  0.1331  0.3263  6.8695

Random effects:
 Groups   Name        Variance Std.Dev.
 Child_ID (Intercept) 7.1121   2.6668
 target   (Intercept) 0.6098   0.7809
Number of obs: 1609, groups:  Child_ID, 68; target, 42

Fixed effects (the reference level is sonority distance 1):
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)         0.38223    0.66659   0.573  0.56636
sonority_distance0 -4.92311    1.15736  -4.254  2.1e-05 ***
sonority_distance4  0.99628    1.04045   0.958  0.33829
sonority_distance2 -0.91196    0.69748  -1.308  0.19104
sonority_distance3  1.61559    0.62385   2.590  0.00961 **
segment_accuracy    0.03667    0.01465   2.503  0.01233 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

VIF score:

```
                      GVIF   Df    GVIF^(1/(2*Df))
sonority_distance   1.128264  4      1.015199
segment_accuracy    1.128264  1      1.062198
```

Test for overdispersion:

```
        chisq         ratio          rdf             p
1115.3356595    0.6966494 1601.0000000    1.0000000
```

C score: 0.94

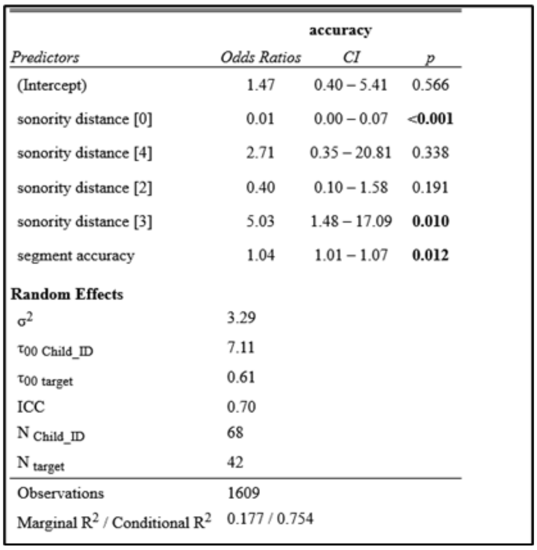
Predicted probabilities of all levels of sonority distance:

```
$`emmeans of sonority_distance`
 sonority_distance  prob     SE      df  asymp.LCL asymp.UCL
 4                  0.8197  0.1375  Inf  0.42339     0.9657
 3                  0.8941  0.0385  Inf  0.79202     0.9493
 0                  0.0121  0.0124  Inf  0.00158     0.0861
 1                  0.6267  0.1563  Inf  0.31180     0.8615
 2                  0.4027  0.1196  Inf  0.20286     0.6412
```

The pairwise comparison of all levels the sonority_distace (not significant pairings are bolded):

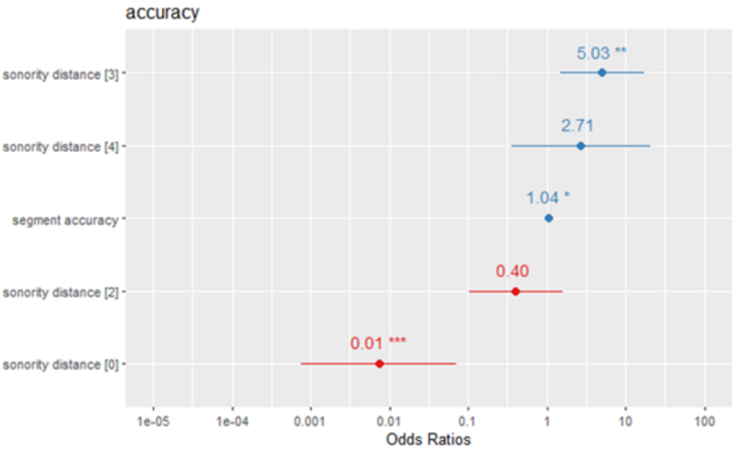
```
$`pairwise differences of sonority_distance`
 1                                         estimate    SE  df  z.ratio p.value
 sonority_distance4 - sonority_distance3   -0.619 0.896 Inf  -0.691  1.0000
 sonority_distance4 - sonority_distance0    5.919 1.329 Inf   4.453  0.0001
 sonority_distance4 - sonority_distance1    0.996 1.040 Inf   0.958  1.0000
 sonority_distance4 - sonority_distance2    1.908 0.952 Inf   2.005  0.4495
 sonority_distance3 - sonority_distance0    6.539 1.037 Inf   6.305  <.0001
 sonority_distance3 - sonority_distance1    1.616 0.624 Inf   2.590  0.0960
 sonority_distance3 - sonority_distance2    2.528 0.455 Inf   5.553  <.0001
 sonority_distance0 - sonority_distance1   -4.923 1.157 Inf  -4.254  0.0002
 sonority_distance0 - sonority_distance2   -4.011 1.069 Inf  -3.752  0.0018
 sonority_distance1 - sonority_distance2    0.912 0.697 Inf   1.308  1.0000
```

Log odds converted into the ratio of odds (the reference level is sonority distance 1):



	accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.47	0.40 – 5.41	0.566
sonority distance [0]	0.01	0.00 – 0.07	**<0.001**
sonority distance [4]	2.71	0.35 – 20.81	0.338
sonority distance [2]	0.40	0.10 – 1.58	0.191
sonority distance [3]	5.03	1.48 – 17.09	**0.010**
segment accuracy	1.04	1.01 – 1.07	**0.012**
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{Child\_ID}}$	7.11		
$\tau_{00\ \text{target}}$	0.61		
ICC	0.70		
$N_{\text{Child\_ID}}$	68		
N_{target}	42		
Observations	1609		
Marginal R^2 / Conditional R^2	0.177 / 0.754		

A plot of main effects:



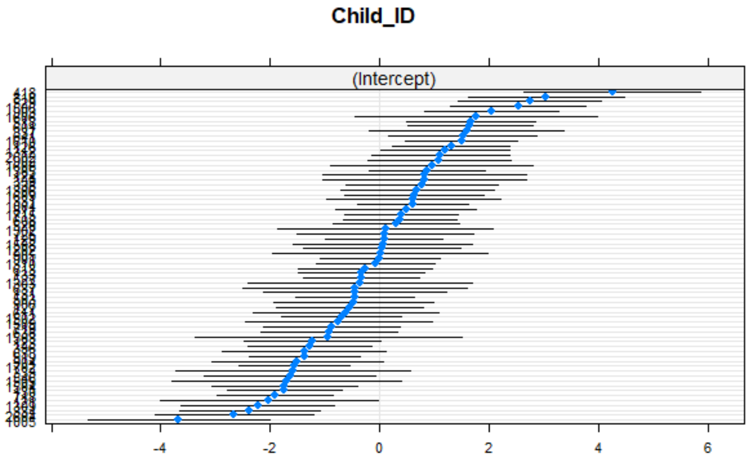
5. The models comparing natural classes of initial clusters

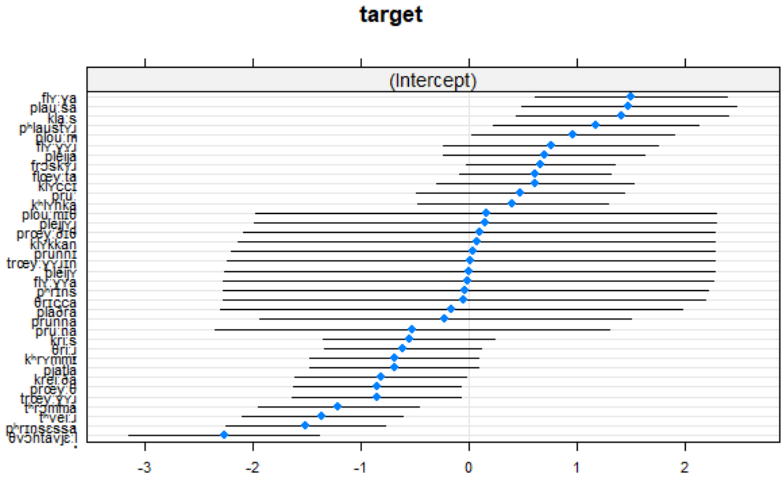
a. Comparison of *stop + approximant* and *fricative + approximant*

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial  ( logit )
Formula: accuracy ~ 1 + segment_accuracy + natural_classes + (1 | Child_ID) +  (1 | target)
   Data: naturalclasses

     AIC      BIC   logLik deviance df.resid
  1023.9   1050.3   -506.9   1013.9     1469

Scaled residuals:
    Min      1Q  Median      3Q     Max
-9.4302 -0.1993  0.1199  0.3214  5.2722

Random effects:
 Groups  Name            Variance Std.Dev.
 Child_ID (Intercept)      2.917    1.708
 target   (Intercept)      1.353    1.163
Number of obs: 1474, groups:  Child_ID, 68; target, 36

Fixed effects:
                                     Estimate Std. Error z value     Pr(>|z|)
(Intercept)                           1.98570    0.36739   5.405      6.49e-08 ***
segment_accuracy_c                    0.17761    0.01999   8.886      < 2e-16 ***
natural_classesfricative_approximant -3.31580    0.59704  -5.554      2.80e-08 ***
```

C score: 0.95

Visualization of the random effects based on participants and tested items:





b. Model comparison a fine-grained distinction of natural classes

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
 Formula: accuracy ~ 1 + segment_accuracy + natural_classes + (1 | Child_ID) +    (1 | target)
   Data: naturalclasses3

    AIC      BIC   logLik deviance df.resid
  944.2    986.2   -464.1    928.2     1398

Scaled residuals:
     Min       1Q   Median       3Q      Max
-11.5993  -0.1654   0.1106   0.3220   5.8717

Random effects:
Groups   Name            Variance Std.Dev.
 Child_ID (Intercept)     3.0550   1.748
 target   (Intercept)     0.1391   0.373
Number of obs: 1406, groups:  Child_ID, 68; target, 35



Fixed effects (stop + lateral is the reference level):
                                   Estimate Std. Error  z value    Pr(>|z|)
(Intercept)                          3.34884   0.36080   9.282    < 2e-16 ***
natural_classesfricative_trill     -4.71897   0.49119  -9.607    < 2e-16 ***
natural_classesfricative_lat       -3.58309   0.45128  -7.940    2.03e-15 ***
natural_classesstop_approximant  -2.38927   0.47124  -5.070    3.97e-07 ***
natural_classesstop_trill          -2.09659   0.33401  -6.277    3.45e-10 ***
segment_accuracy_c                   0.19307   0.02165   8.916    < 2e-16 ***
```

C score: 0.95

```
VIF score:                              GVIF    Df GVIF^(1/(2*Df))
                  natural_classes  1.078461  4        1.009487
             segment_accuracy_c 1.078461  1        1.038490

Test for overdispersion:
chisq               ratio          rdf                    p
1023.9033381    0.7324058 1398.0000000    1.0000000
```

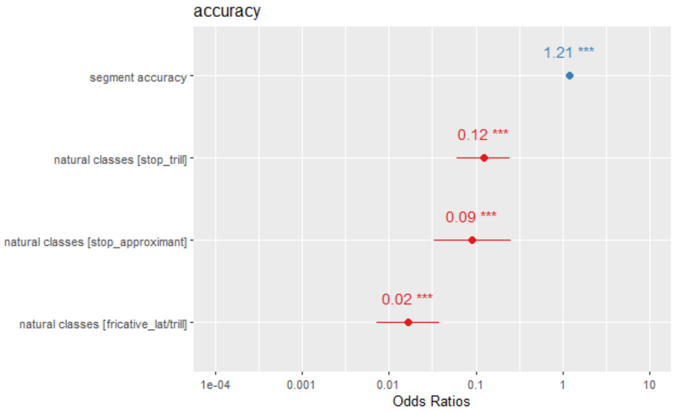
Predicted probabilities of all levels of natural classes:

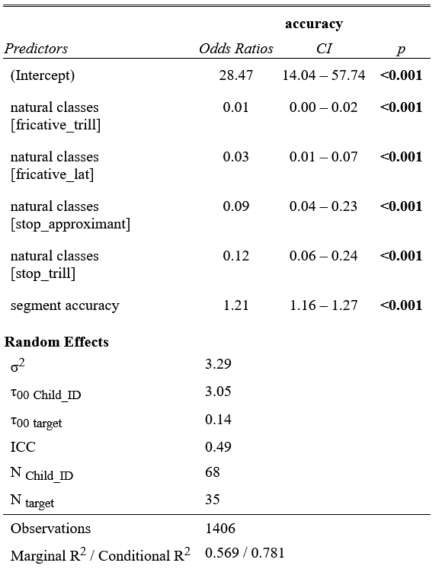
```
$`emmeans of natural_classes`
 natural_classes        prop     SE  df asymp.LCL asymp.UCL
 stop_lateral          0.966   0.361 Inf    2.6417     4.056
 fricative_trill        0.203   0.430 Inf   -2.2120    -0.528
 fricative_lat          0.442   0.404 Inf   -1.0258     0.557
 stop_approximant  0.723    0.443 Inf    0.0922     1.827
 stop_trill              0.778   0.300 Inf    0.6641     1.840
```

The pairwise differences of natural_classes (not significant pairings are bolded):

```
 $`pairwise differences of natural_classes`
 1                                          estimate    SE      df   z.ratio   p.value
 stop_lateral - fricative_trill             4.719    0.491   Inf   9.607    <.0001
 stop_lateral - fricative_lat               3.583    0.451   Inf   7.940    <.0001
 stop_lateral - stop_approximant        2.389    0.471   Inf   5.070    <.0001
 stop_lateral - stop_trill                  2.097   0.334   Inf   6.277    <.0001
 fricative_trill - fricative_lat           -1.136    0.498   Inf  -2.282    0.2249
 fricative_trill - stop_approximant   -2.330   0.535   Inf  -4.354    0.0001
 fricative_trill - stop_trill               -2.622   0.427   Inf  -6.134    <.0001
 fricative_lat - stop_approximant      -1.194   0.507   Inf  -2.354    0.1858
 fricative_lat - stop_trill                 -1.487   0.390   Inf  -3.815    0.0014
 stop_approximant - stop_trill          -0.293  0.423   Inf  -0.692    1.0000
```

Plot of the main effects (stop + lateral is the reference level):





	accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	28.47	14.04 – 57.74	**<0.001**
natural classes [fricative_trill]	0.01	0.00 – 0.02	**<0.001**
natural classes [fricative_lat]	0.03	0.01 – 0.07	**<0.001**
natural classes [stop_approximant]	0.09	0.04 – 0.23	**<0.001**
natural classes [stop_trill]	0.12	0.06 – 0.24	**<0.001**
segment accuracy	1.21	1.16 – 1.27	**<0.001**
Random Effects			
σ^2	3.29		
τ_{00} Child_ID	3.05		
τ_{00} target	0.14		
ICC	0.49		
$N_{\text{Child\_ID}}$	68		
N_{target}	35		
Observations	1406		
Marginal R^2 / Conditional R^2	0.569 / 0.781