

Highly complex syllable structure

A typological and diachronic study

Shelece Easterday

Studies in Laboratory Phonology





Studies in Laboratory Phonology

Chief Editor: Martine Grice

Editors: Doris Mücke, Taehong Cho

In this series:

1. Cangemi, Francesco. Prosodic detail in Neapolitan Italian.
2. Drager, Katie. Linguistic variation, identity construction, and cognition.
3. Roettger, Timo B. Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments.
4. Mücke, Doris. Dynamische Modellierung von Artikulation und prosodischer Struktur: Eine Einführung in die Artikulatorische Phonologie.
5. Bergmann, Pia. Morphologisch komplexe Wörter im Deutschen: Prosodische Struktur und phonetische Realisierung.
6. Feldhausen, Ingo & Fliessbach, Jan & Maria del Mar Vanrell. Methods in prosody: A Romance language perspective.
7. Tilsen, Sam. Syntax with oscillators and energy levels.

ISSN: 2363-5576

Highly complex syllable structure

A typological and diachronic study

Shelece Easterday


Easterday, Shelece. 2019. *Highly complex syllable structure: A typological and diachronic study* (Studies in Laboratory Phonology). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/000>

© 2019, Shelece Easterday

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: no digital ISBN

no print ISBNs!

ISSN: 2363-5576

no DOI

ID not assigned!

Cover and concept of design: Ulrike Harbort

Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

For Elise, Ada, Astrid, and Maria:

May you always be curious!

Contents

Acknowledgments	iii
Abbreviations	v
1 Syllable structure patterns in sample	1
1.1 Introduction	1
1.1.1 Crosslinguistic studies of syllable structure	1
1.1.2 Considerations in the current chapter	4
1.2 Methodology	6
1.2.1 Patterns considered	6
1.2.2 Status of inserted vowels	9
1.2.3 Edges of categories	13
1.2.4 Coding	15
1.3 Results for language sample	16
1.3.1 Maximal onset and coda sizes	17
1.3.2 Relationship between onset and coda complexity	18
1.3.3 Syllable structure complexity and obligatoriness of syllable margins	19
1.3.4 Vocalic nucleus patterns	21
1.3.5 Syllabic consonants	23
1.3.6 Morphological patterns	27
1.4 Properties of highly complex syllable structure	34
1.4.1 Examples of Highly Complex syllable patterns in sample	34
1.4.2 Prevalence of Highly Complex syllable patterns within languages	37
1.4.3 Acoustic and perceptual characteristics	44
1.5 Discussion	47
References	53
Index	59
Name index	59

Contents

Language index 59

Subject index 59

Acknowledgments

This book is a modified version of my Ph.D. dissertation, which was defended in May 2017 at University of New Mexico. While the book features an updated language sample, the findings and interpretations are largely unchanged from the original study. Neither work would have been possible without the mentorship and support of many people along the way.

I am particularly grateful to the members of my dissertation committee. My advisor, Caroline Smith, provided helpful and constructive feedback on all aspects of the dissertation and my progress through the Ph.D. Joan Bybee and Ian Maddieson introduced me to language change and phonological typology, respectively, and I have greatly benefited from the inspiring work and mentorship of both. I thank Ioana Chitoran for many thought-provoking discussions of the findings and the next steps in this line of research. Thanks also to Bill Croft.

While writing the original dissertation I was fortunate to have the financial support of the Joseph H. Greenberg Endowed Fellowship (2011–2015) and the Russell J. and Dorothy S. Bilinski Fellowship (2016–2017).

Since October 2017, I have been employed as a postdoctoral researcher at Laboratoire Dynamique du Langage (UMR5596, CNRS & Université Lyon 2) in Lyon, France. While revising this book, I have been funded by LABEX ASLAN (ANR-10-LABX-0081) of the Université de Lyon within the French program Investissements d’Avenir. DDL has been a warm, welcoming, and intellectually exciting environment in which to work, and for this I am grateful to my supervisor François Pellegrino, laboratory director Antoine Guillaume, and Anetta Kopecka, the head of the Description, Typologie, et Terrain (DTT) research axis. Special thanks also go to my fellow “postdoc ladies” – Natalia Chousou-Polydouri, Marine Vuillermet, Laetitia de Almeida, Elisa Demuru, and Natasha Aralova – for their friendship and moral support during this stage of our lives and careers, and to DDL alumna Kasia Janic for the use of her magical apartment.

Many specialists kindly answered my inquiries about details of the languages and language families discussed in this book, including Doug Marmion (Wutung), Matthew Carter (Bashkir), Netra P. Paudyal (Darai), Jonathan Bobaljik (Itelmen), Andrey Filchenko (Eastern Khanty), Bonny Sands (Hadza), Kirk Miller (Hadza),

Acknowledgments

Logan Sutton (Kiowa-Tanoan), Rosa Vallejos (Cocama-Cocamilla), Sylvia Tufvesson (Semai), Kasia Wojtylak (Murui Huitoto), and Albert Alvarez Gonzalez (Uto-Aztecan). Tim Zingler, Ricardo Napoleão de Souza, Lexie Adams, and Marc Adams helped me with the translation of reference material from the original German, Portuguese, and Spanish.

I am grateful to many dear friends for their support over the years that this work was in progress, including Caron Treloar, Lexie Adams, Erin Debenport, Cora McKenna, Francesca Di Garbo, Giorgio Iemmolo, Logan Sutton, Jason Timm, Ricardo Napoleão de Souza, Fabian Armijo, Nick Mathews, Amanda Delgado Galvan, and Sylvia Tufvesson.

Finally, to my family: thank you for everything.

Abbreviations

1	first person
3	third person
AGT	agentive
CAUS	causative
CL	noun class
CONT.P	past continuous
DAT	dative
DEM	demonstrative
DUB	dubitive
F	feminine
FUT	future
INCL	inclusive
INF	infinitive
INTERROG	interrogative
ITER	iterative
M	masculine
NMLZ	nominalizer
NOM	nominative
PL	plural
POSS	possessive
PS	predicate specifier
PST	past
REALIS	realis
RECP	reciprocal
REFL	reflexive
REL	relativizer
SG	singular
RNS	translated from the original by Ricardo Napoleão de Souza
SME	translated from the original by Shelece Easterday
TZ	translated from the original by Tim Zingler

1 Syllable structure patterns in sample

In this chapter properties of syllable structure in the language sample are presented. In §1.1 common topics of research in crosslinguistic studies of syllable structure and specific considerations in the current study are discussed. In §1.2 the methodology and coding strategies are described. In §1.3 the results on onset, coda, and nucleus patterns, as well as morphological constituency patterns in maximal clusters, are presented for the language sample as a whole. In §1.4 a more detailed analysis of the syllable patterns of the languages in the sample with Highly Complex syllable structure is presented. In §1.5 the findings are summarized and related to following chapters.

1.1 Introduction

1.1.1 Crosslinguistic studies of syllable structure

A common approach to studying syllable structure on a crosslinguistic scale is to compare canonical syllable patterns across languages. This is the range of occurring syllable patterns in a language represented as a sequence of Cs for consonants and Vs for vowels, with parentheses indicating optional components of the onset, nucleus, and coda (e.g., (C)(C)(C)V(C)(C)(C)(C) for English). Many databases of phonological patterns include canonical syllable structure as one of the coded features along with consonant and vowel phoneme inventories; e.g., the World Phonotactics Database (Donohue et al. 2013), LAPSyD (Maddieson et al. 2013), and a modified version of the WALS 100-language sample presented in Gordon (2016). Though it is a very general measure, the size and shape of canonical syllable patterns can be used to categorize languages in such a way as to capture predominant global trends (cf. Maddieson 2013).

More often, crosslinguistic studies of syllable patterns are focused on finer-grained aspects of syllable structure, including sub-syllabic constituents. A number of studies investigating the properties of syllable margins have revealed trends regarding the size, voicing, place, manner, and sonority of consonant sequences in the onset and coda, as well as implications regarding the makeup of onset and coda inventories (cf. Greenberg 1965, see Chapter ??). Large-scale studies

1 Syllable structure patterns in sample

of syllable margins are typically limited to biconsonantal clusters, as these are crosslinguistically the most frequent cluster type, but larger clusters have been explored as well (VanDam 2004 , and a few of the analyses in Greenberg 1965). A few typological studies of cluster patterns focus specifically on implicational relationships in obstruent clusters. For instance, Morelli (1999 ; 2003) examines implicational relationships in biconsonantal onsets composed of stops and fricatives in 30 languages, while Kreitman (2008) reports implicational relationships among obstruent clusters of mixed sonority and voicing in 62 languages. There are also studies which examine the patterns of simple onsets and codas: Rousset (2004) compares, among other things, the relative proportion of consonants which occur in onset versus coda position in 15 diverse languages. Nucleus patterns have also been the subject of typological investigation, with much of the research emphasizing the sonority-based implications that can be gleaned from global distributions of syllable nuclei patterns (Blevins 1995 ; Zec 2007). As discussed in §??, Bell (1978) explores global and areal patterns of syllabic consonants in a sample of 182 diverse languages, deriving implicational generalizations regarding the manner and place of articulation of such sounds. Hoard (1978) is a survey of syllabic obstruent patterns in languages of the Pacific Northwest and Northwest Plateau regions of North America, and includes analyses of data from five diverse languages as well as brief references to others.

Relationships among the different constituents of the syllable have also been investigated crosslinguistically. An analysis of syllable frequencies in the lexicons of five languages revealed that simple onsets and nuclei — CV sequences — may combine relatively freely (Maddieson & Precoda 1992). By contrast, relationships between other pairs of subsyllabic constituents tend to show more restrictions across languages. Onsets and codas have often been treated independently of one another with respect to their internal patterns and relationships to the larger syllable; however, Davis & Baertsch (2011) show that the coda and the second member of a biconsonantal onset are often restricted to the same subset of consonants in a language. Similarly, Blevins (2006) observes that languages with only open syllables tend to have optional onsets. Gordon (2006) investigates the wide range of behavior exhibited by languages with respect to issues of syllable weight, which may be dependent upon complex combinations of nucleus and coda patterns, and minimal requirements for root and/or word structures, which may implicate onset patterns in addition to those of the rime.

Many researchers have investigated the properties of syllables in contact with one another or within the context of larger domains. Crosslinguistically, syllabification patterns are such that sonority tends to fall from the coda of one syllable

ble to the onset of the next; this observation is supported by evidence from diachronic processes of sound change in various languages (Hooper 1976 ; Murray & Vennemann 1983). Syllable margins can also exhibit differing patterns according to their position within larger domains. Côté (2011) observes crosslinguistic asymmetries in coda patterns in final versus medial position in stem, word, and phrasal domains. Similarly, ‘prominent’ environments such as word- and phrase-initial syllables and stressed syllables are often the locus of the highest amount of phonemic contrast and variety in syllable margins in a language, though there are areal exceptions to this global trend. For example, Gasser & Bowerman (2014) note more onset restrictions in word-initial as compared to word-internal position in Australian languages.

Crosslinguistic studies also investigate the distribution of syllable types within syllable inventories, lexicons, and words of varying lengths in order to determine frequency patterns and uncover implicational hierarchies. The presence of VC structures in a language, for example, generally implies the presence of V, CV, and CVC structures as well (Blevins 1995). Frequency distributions of syllable types in the lexicons of diverse languages reveal a heavy crosslinguistic dominance of the CV type, despite wide variation in canonical syllable structures (Rousset 2004 ; Vallée et al. 2009). There is also evidence of a relationship between syllable length and word length. -Fenk & Fenk-Oczlon (1993) tested Menzerath’s Law (paraphrased as ‘the bigger the whole, the smaller the parts’) in a sample of 29 languages and found a significant negative linear correlation between the number of syllables per word and the number of phonemes per syllable.

A facet of syllable structure that has received limited crosslinguistic treatment is that of the morphological constituency of clusters at syllable margins. Several of the studies listed above (e.g., Greenberg 1965 , Morelli 1999 ; Kreitman 2008) exclude morphologically complex clusters from analysis on the assumption that these clusters may exhibit different patterns than clusters found within the boundaries of a single morpheme. Dressler & Dziubalska-Kołaczyk (2006) examine syllable patterns in a sample of four Indo-European languages and propose two types of clusters defined by morphological constituency: phonotactic clusters, which are phonologically motivated and morpheme-internal, and morphonotactic clusters, which come about through morphological processes. They conclude that the latter are often longer and more phonologically marked in comparison to the former (see also Dressler et al. 2010 ; Orzechowska 2012 , and Dressler et al. 2015 for other studies in this vein).

1.1.2 Considerations in the current chapter

While all of the above issues are important in developing a detailed understanding of crosslinguistic syllable patterns, for practical purposes the current study is limited to exploring just a few of these in depth. In this chapter I investigate issues of syllable structure that are directly pertinent to addressing the main research questions and hypotheses of this study. Specifically, I limit the scope of analysis here to features of syllable structure which have been previously demonstrated or hinted in the literature to be correlated with other linguistic, and especially phonological, features, and those which I hypothesize may reveal clues about the diachronic development of highly complex syllable structure. Additionally, I examine the syllable patterns of the languages in the Highly Complex portion of the sample in greater detail than the other categories, in order to develop a better understanding of the coherence of that group. The features under consideration are listed in (1)-(2):

- (1) **Features examined for entire sample**
 - a. *Canonical syllable structure*
 - b. *Nucleus patterns*
 - c. *Morphological constituency of maximal syllable margins and syllabic consonants*
- (2) **Features examined for languages in the Highly Complex category**
 - a. *Specific Highly Complex syllable patterns occurring*
 - b. *Restrictions on Highly Complex syllable patterns*
 - c. *Relative frequency of Highly Complex patterns within languages*
 - d. *Phonetic characteristics of Highly Complex clusters*

The coding of feature (1a), canonical syllable structure, is motivated by previous findings in the research. As mentioned in Chapter ??, positive correlations have been established between syllable structure complexity (defined categorically and holistically with reference to canonical syllable structure) and consonant phoneme inventory size (Maddieson 2013). A positive relationship has also been established between syllable structure complexity and the number of consonants belonging to certain classes within a language (Maddieson et al. 2013). Gordon (2016) has demonstrated that the trend by which consonant phoneme inventory size increases with syllable complexity occurs on a more fine-grained level as well, when complexity is measured as the combined sum of canonical onset and coda constituents. Along another line of inquiry, Blevins (2006 : 336)

has described a crosslinguistic tendency by which languages without codas tend to have optional onsets. This suggests a relationship between obligatoriness of constituents and canonical syllable structure patterns which, to my knowledge, has not been investigated in a language sample controlled for syllable structure complexity.

Nucleus patterns (1b), and syllabic consonant patterns more specifically, have also been suggested in the literature to bear some relation to syllable structure complexity. On the one hand, the holistic phonological typology proposed by Isačenko (1939) predicts that ‘vocalic’ languages — those which have shorter consonant sequences and relatively higher vowel/consonant ratios — are more likely to develop syllabic consonants than ‘consonantal’ languages. This suggests that we might expect a greater prevalence of syllabic consonants in languages of the Simple and Moderately Complex categories. On the other hand, Bell (1978) notes that syllabic consonants often come about through vowel deletion, often in unstressed syllables, which is also known to be a common source of consonant clusters. Based on that observation, we might expect a higher occurrence of syllabic consonants in languages with Complex or Highly Complex syllable structure. The way that the language sample is constructed allows us to test these predictions directly. My hypothesis, rooted in the discussion of speech rhythm typology in Chapter ??, is that the latter prediction will be borne out (3).

- (3) *Languages with more complex syllable structure are more likely to have syllabic consonants.*

Feature (1c), morphological constituency, is related to both of the issues discussed above. Greenberg (1965 : 250) and Dressler & Dziubalska-Kołaczyk (2006) predict that as the size of a syllable margin increases, so does the probability that it contains morpheme boundaries. Additionally, syllabic consonants are often noted to be largely or entirely restricted to grammatical particles and affixes (Bell 1978 : 159), suggesting additional potential interactions between syllable structure complexity and morphology if syllabic consonants are found to occur more frequently in languages on one end of the syllable complexity scale. In coding for (1c) I consider the morphological patterns of the largest syllable margins in each language, as well as the kinds of morphemes in which syllabic consonants occur in the languages which have them. These analyses will provide a rough measure of the relationship between morphology and the syllable patterns of a language, which can then be explored in greater depth at a later point, especially if the results point towards a heavy role of morphology in the development of highly complex syllable structure. My hypotheses regarding morphological con-

page
numbers
refer to
which
edition?

1 Syllable structure patterns in sample

stitutency follow (4a-b):

- (4) a. *As syllable structure complexity increases, so does the likelihood that the largest syllable margin types in a language will be morphologically complex.*
- b. *As syllable structure complexity increases, so does the likelihood that syllabic consonants occurring in a language will belong to grammatical elements.*

The examination of features (2a-d) is intended to develop a more detailed picture of the syllable patterns of the languages in the Highly Complex portion of the sample. The purpose of this portion of the analysis is twofold: to determine whether the languages form a coherent phonological type with respect to their syllable patterns, and to gather information that may be relevant to uncovering how these structures come about over time. The specific syllable patterns falling under the definition of Highly Complex will be examined for each language and compared with those of the other languages of the group to determine how similar they are to one another. In an attempt to characterize the prevalence of Highly Complex syllable patterns within each language, restrictions on the patterns will be examined for each language, and information gathered on the relative frequency of the patterns within the language. Finally, descriptions of the phonetic characteristics of Highly Complex syllable structures will be noted and compared.

Apart from those which are aimed at testing the three specific hypotheses listed above, the analyses in this chapter are largely exploratory in nature. The findings here are intended to provide a baseline characterization of the syllable patterns of the four categories of the language sample. Additionally, the in-depth examination of the Highly Complex portion of the sample will serve as a point of reference in which to ground the results of later chapters.

1.2 Methodology

1.2.1 Patterns considered

As discussed in Chapter ??, interpretations of syllable patterns may vary dramatically according to the theoretical framework which is used. For example, even simple coda consonants and biconsonantal onsets, both crosslinguistically common patterns, are problematic in a strict CV approach, which must posit empty nuclei for phonetic occurrences of these structures (Lowenstamm 1996). Though

less restrictive models of the syllable may accept small onset and coda clusters, large clusters and syllabic obstruents are typically considered problematic, with members being relegated to an appendix or left entirely unsyllabified at the level of phonological representation (Vaux & Wolfe 2009 ; Bagemihl 1991). In language documentation and description, authors often choose a model which is compatible with the occurring syllable patterns of the language. However, some researchers work within a strict theoretical framework in which the abstract syllable structure has been predetermined, and the patterns of the language are made to fit into the model. This causes obvious complications for crosslinguistic comparison of syllable patterns.

Because such a wide range of frameworks is used by researchers in their descriptions of syllable structure, and because the choice of model may have strong implications for the patterns reported, I opt instead to use a definition of canonical syllable structure motivated largely by the invariant phonetic patterns observed at word margins. Here onset clusters include patterns occurring word-initially before a vowel (and, if they differ at all from word-initial patterns, clusters occurring word-internally before a vowel where syllabification as an onset is supported by language-internal evidence). Likewise, coda clusters include patterns occurring word-finally after a vowel (with the same considerations given to word-internal post-vocalic patterns). In this view, all consonants are syllabified as part of an onset or coda, or, in the case of syllabic consonants, as a nucleus. Morphologically complex pre-/post-vocalic sequences, such as the word-initial biconsonantal sequence in (5) below, are thus considered as onset or coda patterns, in addition to morpheme-internal sequences.

- (5) **Tzeltal** (*Mayan*; Mexico)
 /s-ʈʰuht/
 POSS.3-stomach
 ‘his stomach’ (Polian 2006 : 24)

Pre-/post-vocalic sequences in phonological words are also considered as syllable margins, even if they belong to separate units syntactically or orthographically (6).

- (6) **Polish** (*Indo-European*; Poland)
 z pstrągiem
 /spstrɔŋjem/
 ‘with (the) trout’ (Jassem 2003 : 103)

1 Syllable structure patterns in sample

Syllable margins reported to occur only in recent loanwords or as a result of variable phonetic processes, such as vowel elision in rapid speech, were not considered in the determination of canonical patterns, or included in the present analysis (the latter issue will be treated extensively in Chapter ??, which examines vowel reduction in the sample). Here ‘canonical’ is used in the sense of syllable patterns which are regular and not reported to occur in variation with other patterns. For instance, /tk/ would not be characterized as a canonical onset in American English, because it is an optional, if for some words ubiquitous (Napoleão de Souza 2019), variant of a form which preserves the original vowel: *tequila* [tʰəkʰilə] ~ [tʰkʰilə].¹ Returning to the issue of abstract models of the syllable versus observed patterns, when syllable margins were described in one way at the phonological level but reported as consistently exhibiting a different pattern at the ‘surface’ level, the latter was taken to be representative of syllable patterns in the language. For example, sequences analyzed as clusters at an abstract level but which were invariably split by phonological epenthesis, as illustrated by the Maybrat example below, were not considered to be canonical clusters (example 7; see the following section for a further discussion of epenthesis).

(7) **Maybrat** (*Maybrat-Koron*; Indonesia)

An epenthetic vowel invariably occurs between two consonants in word-initial position.

/pnem/

[panem]

‘it is flat’ (Dol 2007 : 35-6)

Sounds involving multiple articulations or which consist of phonetic sequences, such as labialized consonants, prenasalized stops, or diphthongs, present complications for determining canonical syllable patterns, which are traditionally expressed in segmental terms. Wherever the author has presented convincing language-internal evidence for the phonological unity of such sounds, I have considered them to be single segments for the purpose of syllable structure coding. The issue of complex segments will be discussed in greater detail in Chapter ??, where I analyze the segmental inventory patterns in the language sample.

As described in §??, similar patterns observed in languages with large tautosyllabic obstruent clusters and those with syllabic obstruents led me to include the latter in the Highly Complex category, so long as syllabic obstruents were

¹Moreover, in this case the much rarer pronunciation of [tʰəkʰilə] is still considered an acceptable variant by speakers.

observed to result in word-marginal sequences of three obstruents or longer in those languages. However, syllabic obstruents were not explicitly coded as part of the onset or coda. Instead, Highly Complex patterns involving long tautosyllabic obstruent clusters and those involving syllabic obstruents were coded separately so that these patterns could be disambiguated if necessary in later analyses. For languages with syllabic obstruents, the author's description of 'true' tautosyllabic onset and coda patterns was taken to be canonical. Separately, the largest word-marginal consonant sequences occurring as a result of syllabic obstruents in those languages were coded as such but not considered to be maximal coda or onset structures.

It is important to note that while the presence of syllabic obstruents was a partial criterion in determining whether a language has Highly Complex syllable structure, the presence of syllabic nasals and liquids (where liquid is defined as a trill, tap/flap, or lateral approximant) was never used as a diagnostic for membership in this or any category. The reasoning behind this is related to how these patterns are described and analyzed in the language references. Recall that syllabic obstruents and large tautosyllabic obstruent clusters often co-occur in languages, and that it is also common for the same language to be analyzed as having either pattern, to the exclusion of the other, by different authors. This was the case for a number of languages in the sample, including Chipaya, Nimboran, Tashlhiyt, and Yine: all of these languages have been reported in some descriptions to have syllabic obstruents, and in others to have larger obstruent clusters. By comparison, only two languages were described as having syllabic nasals or liquids as an alternative analysis to larger tautosyllabic clusters (Georgian and Yine). This is despite the fact that syllabic nasals and liquids are much more commonly reported in the sample than syllabic obstruents (see §1.3.5). Nasals and liquids do not seem to be as susceptible as obstruents to ambiguous or competing interpretations with respect to syllabicity. This is perhaps not surprising, given the perceptual properties of nasals and liquids. The reported patterns justify separate treatment of syllabic nasals and liquids on one hand, and syllabic obstruents on the other, in the coding of canonical syllable patterns in the sample.

1.2.2 Status of inserted vowels

The Maybrat example in (8) brings up another important issue in the analysis of syllable structure, which is the status of inserted vowels. Compare the Maybrat example, reproduced in (8), with the examples from Camsá (9a-b).

- (8) **Maybrat** (*Maybrat-Karon*; Indonesia)

1 Syllable structure patterns in sample

An epenthetic vowel invariably occurs between two consonants in word-initial position.

/pnem/

[panem]

‘it is flat’ (Dol 2007 : 35-6)

(9) Camsá (isolate; Colombia)

“A nonphonemic transitional voicoid [ə] occurs between stop plus stop [...] Initial fricatives [...] have optional off-glide before nonfricative consonants at a different place of articulation.”

a. /tkanipe/

[t^əkanipe]

‘broken’

b. /ftʂɛŋga/

[f^uʂɛŋga]

‘black’ (Howard 1967 : 81)

There are several important differences between the descriptions of the Maybrat and Camsá examples. In the Maybrat example the vowel is described as epenthetic, while in the Camsá example it is described as a nonphonemic transitional vocoid or an off-glide. In Maybrat, the process is described as invariable, while in Camsá it is described, in the case of the off-glide, as optional. Finally, in Maybrat, the epenthetic vowel is transcribed as [a], while in Camsá it is transcribed as superscripted [ə] in one case and [u] in another. These differences in the descriptions and transcription conventions suggest different phenomena. This raises the question of whether the patterns in Maybrat and Camsá should be treated separately for the purposes of syllable structure analysis.

In a typological survey of reported inserted vowel patterns, Hall (2006) makes a distinction between epenthetic vowels and intrusive vowels. She argues that the motivations for the two types of inserted vowels are quite different. The function of epenthetic vowels is to provide a nucleus to repair marked or non-occurring syllable structures in a language; that is, it has the effect of producing syllable patterns that are already attested in the language. A textbook example of this kind of process is the adaptation of loanwords from one language into another; e.g. English *technostress* is borrowed into Japanese as *tekunosutoresu* because the consonant clusters and codas in the original form are not part of the sound pattern of the borrowing language (Kay 1995 : 69). Hall states that epenthetic vowels tend to be ‘visible’ to phonological processes, behaving like

syllable nuclei for the purposes of stress assignment and other processes, and speakers are generally aware of their presence. For example, in Mono, epenthetic vowels occur to ‘repair’ monomoraic lexical words: /ʒĩ / > [ĩ ʒĩ] ‘tooth’ (Hall 2006 : 6, citing Olson 2003).

By contrast, intrusive vowels are not structurally motivated but come about through natural processes of gestural retiming and overlap, and are simply an acoustic effect of these transitions. For example, in Kekchi, an intrusive vowel identical in quality to the preceding vowel may appear within final clusters consisting of a glottal stop followed by a consonant: /poʔt/ > [poʔot] ‘blouse’ (Hall 2006 : 7, citing Campbell 1974). This is analyzed not as the addition of a new vowel articulation, but as the offset of the vowel gesture carrying over the duration of the glottal stop articulation, which does not require an oral articulation, such that the effects of the vowel can still be heard between the glottal stop and the onset of the following consonant articulation. Hall states that intrusive vowels tend to be ‘invisible’ to such phonological processes, and speakers are typically unaware of their presence, or if they are aware, view them as optional.

The properties described above — behavior with respect to phonological processes and speaker awareness — are for Hall the primary means of determining whether an inserted vowel is epenthetic or intrusive. She argues persuasively for the use of speaker intuition in making these determinations, giving several examples of cases in which the phonological behavior of a vocalic element matches the native speaker’s, rather than the fieldworker’s, intuitions about syllabicity. This is supported by growing evidence in the experimental literature that listeners are biased by the timing patterns of their own language when identifying the syllable patterns of another language (Kwon et al. 2017). However, while information regarding native speaker intuition is valuable, it is often unreported. Based on the typological patterns represented in her survey, Hall develops a set of additional features which tend to be associated with epenthesis and intrusion. Epenthetic vowels tend to have the following characteristics: (i) the vowel quality is fixed or copied from a neighboring vowel; (ii) the vowel occurs regardless of speech rate; and (iii) the vowel repairs a structure that is likely to be avoided through other processes in the same language. Intrusive vowels tend to have these characteristics: (i) the vowel quality is neutral, or influenced by the place of articulation of the surrounding consonants; (ii) the vowel is typically found in heterorganic clusters; (iii) the vowel is often optional, has variable duration and voicing, and may disappear as speech rate increases; and (iv) the vowel does not seem to have a repairing function (Hall 2006 : 391).²

²As noted by Browman & Goldstein (1992 : 53), epenthetic vowels may have their origin in the

1 Syllable structure patterns in sample

Returning to (8)-(9) above, we see that the Maybrat example falls under Hall's definition of epenthetic vowels, while the Camsá example falls under her definition of intrusive vowels. Further descriptions by the authors support this view. Dol (2007 : 36) presents instrumental evidence showing that epenthetic [a] is as prominent as other vowels in the language. Meanwhile, examples given by Howard indicate that the transitional vocoid is not counted for the purposes of stress assignment. This suggests that the Maybrat example [panem] is best analyzed as a CV.CVC structure, while the Camsá examples [tʰkaniŋe] and [fʰʈʂɛŋga] are best analyzed as having initial onset clusters.

The matter of vowel intrusion is relevant to the issue of syllabic consonants as well. Recall the Cocopa example discussed in §??, reproduced below as (10).

(10) **Cocopa** (*Cochimi-Yuman*; USA and Mexico)

check
diacritics

/pʰʃxmuka' p/

[pʰ.ʃx^a.mu.kə' p]

'he embraced her' (Crawford Jr 1966 : 43)

Crawford proposes that some unstressed syllables in Cocopa may be entirely consonantal, consisting only of an onset or onset and coda, but with a predictable 'murmur' vowel functioning as a phonetic peak in such cases (1966: 34). What is described as the murmur vowel here has properties of an intrusive vowel: its quality is determined by that of surrounding consonants, and it is transcribed as a superscript, suggesting brevity or an offglide status.

Similar descriptions of other languages in the sample suggest an association between intrusive or transitional elements and consonants analyzed as syllabic (e.g., in certain environments in Tashlhiyt; Dell & Elmedlaoui 2002). Hargus & Beavert (2006) list several languages in which arguments for consonant (specifically obstruent) syllabicity are rooted in the distribution of 'epenthetic' vowels. Additionally, Bell (1978 : 185-6) reports that speakers are unaware of the short vocalic elements associated with syllabic obstruents in Koryak, similar to reports of speaker intuition of intrusive vowels in consonant clusters. Hall notes crosslinguistic associations between vowel intrusion, aspiration (also analyzable as voiceless vowel intrusion), and syllabic consonants, and suggests that all of these phenomena are different acoustic manifestations of fundamentally similar processes of gestural overlap (2006: 413). This issue will be revisited in §1.4.3, when I dis-

intrusive elements arising from gestural organization. Hall gives several examples of historical cases of vowel epenthesis that may have started with vowel intrusion and then phonologized; e.g., Irish Gaelic *gorm* > *gorəm* 'blue' (2006: 35).

cuss the phonetic properties of languages in the Highly Complex portion of the sample.

In coding for syllable structure patterns in the sample, wherever possible I used the diagnostics proposed in Hall (2006) to determine the status of reported inserted vowels.

1.2.3 Edges of categories

After syllable patterns in the sample were determined with reference to the above criteria, these patterns were recorded and coded for analysis. First, the syllable patterns were used to classify the languages into the four categories of syllable structure complexity as defined in Maddieson (2013) and §??, and reproduced here:

Simple: languages in which the onset is maximally one C, and codas do not occur.

Moderately Complex: languages in which the onset is maximally two Cs, the second of which is a liquid or a glide; and/or the coda consists of maximally one C.

Complex: languages in which the maximal onset is two Cs, the second of which is a C other than a liquid or a glide, or three Cs, so long as all three are not obstruents; and/or the maximal coda consists of two Cs, or three Cs so long as all three are not obstruents.

Highly Complex: languages in which the maximal onset or coda consists of three obstruents, or four or more Cs of any kind; and/or in which syllabic obstruents occur, resulting in word-marginal sequences of three or more obstruents.

A common complication for typological work is that languages sometimes do not fall neatly into the categories defined by the researcher. Like languages themselves, linguistic features are dynamic and constantly changing form, however slowly or subtly. There is often some ambiguity at play when categorizing languages according to some structural feature, because a language may exhibit behavior characteristic of several categories. In this study I took advantage of such ambiguity in constructing the language sample in order to increase genealogical diversity or the representation of syllable structure complexity categories in some regions.

1 Syllable structure patterns in sample

The Simple syllable structure pattern is crosslinguistically rare, so in some cases languages were admitted to the category despite having minor exceptions to canonical (C)V patterns. For example, several languages were categorized as having Simple syllable structure despite being reported to have codas in a small handful of lexical items (e.g., five words in Cavineña, [Guillaume 2008](#) : 31; two in Saaroa, [Pan 2012](#) : 32f). Two of the North American languages included in this category, Towa and Ute, can be argued not to have Simple syllable structure in the strictest definition of the term. In Towa, there are no complex onsets. Simple codas /f/ and /l/ occur in the language, but are rarely produced as such by speakers. For instance, word-final /f/, which always corresponds to Inverse suffix /-f/, is omitted from a phrase-medial noun stem unless it is followed by a vowel-initial pronominal prefix, in which case it is resyllabified as an onset ([Yumitani 1998](#) : 22-24). Coda /f/ may occur utterance-finally, but because the language is predominantly verb-final and this suffix attaches to nouns, the frequency of this pattern is extremely rare in natural discourse (Logan Sutton, p.c.).³ In Ute, syllable structure is almost entirely of the shape CV(V), but a recent process devoicing unstressed vowels in the language has resulted in some invariant codas and patterns that could be analyzed as sequences of oral consonants and glottal fricatives ([Givón 2011](#) : 20-3, 28). It would appear that Ute has until recently had Simple syllable structure, but is in the process of rapidly developing more complex syllable patterns. Though neither Towa nor Ute present uncontroversial cases of languages with Simple syllable structure, I judge their patterns to be close enough to justify their inclusion as such, thereby increasing the representation of North American languages in this category.

On similar grounds, Eastern Khanty was admitted to the Moderately Complex category. The syllable patterns of this language include occasional coda clusters which are always a result of derivation or inflection in the language. Typically when this happens, vowel epenthesis is employed “robustly and productively” such that most of these sequences are not realized as clusters ([Filchenko 2007](#) : 55). However, derived coda clusters with a sonorant preceding a homorganic stop (e.g., *lol-t* ‘crack, dent-PL’) are sometimes retained as such. The author states that the probability of such consonant clusters occurring is extremely low. As a complex coda, this pattern falls under the definition of Complex syllable structure, but since its status in the language is extremely marginal, I place this language into the Moderately Complex category.

Yine was admitted to the Highly Complex category despite being a some-

³In a 244-word text included in [Yumitani \(1998\)](#), there are no examples of phonetic codas in Towa.

what ambiguous case. Two major descriptions of Yine (Hanson 2010 ; Matteson 1965) describe the occurrence of biconsonantal and triconsonantal onset clusters in the language. Both describe onset clusters as being relatively unrestricted: “[c]onsonant clusters in Yine show enough range in attested combinations, both word-initially and word-internally, to suggest that there are no sonority-based restrictions imposed on them” (Hanson 2010 : 27). Examples of biconsonantal onsets are plentiful and include combinations as varied as /tʃk/, /sp/, and /ns/. Matteson (1965 : 24) gives few examples of triconsonantal onsets and states that the very low frequency of these shapes had decreased in comparison to a count made a decade previously. However, Hanson, writing 45 years later, writes that “words beginning with three consonants in a sequence are very common” (2010: 27). She describes these clusters as resulting from the affixation of a Class 2 pronominal prefix (/n-/ , /p-/ , /t-/ , /w-/ , /h-/) to a stem beginning with a biconsonantal cluster; e.g. /p-knoya-te/ ‘your tortoise’ (2010: 26). Other examples given include /pɕr/, /nmʃj/, and /nʃjk/, but there are no explicit examples of three-obstruent clusters listed. Whether this is due to non-occurrence is unclear, and the conflicting descriptions of the frequency of triconsonantal onsets does not help shed light on the issue. Nevertheless, I include this language in the Highly Complex category to increase representation of the category in the South American portion of the sample, acknowledging that it may be a very marginal case.

1.2.4 Coding

The analyses described in §1.2.1-3.2.3 were conducted as part of the process of constructing the language sample (Chapter ??). After the sample was constructed, more specific information on syllable structure was collected for each language, and coded as described below:

Size of maximal onset: in number of Cs

Size of maximal coda: in number of Cs

Onset obligatory: Yes, No

Coda obligatory: Yes, No

Vocalic nucleus patterns: Short vowels, Long vowels, Diphthongs, Vowel sequences

Syllabic consonant patterns: Nasal, Liquid, or Obstruent

Size of maximal word-marginal sequences with syllabic obstruents: in number of Cs

1 Syllable structure patterns in sample

Predictability of syllabic consonants: Phonemic, Predictable from word/consonantal context, or Varies with CV sequence

Morphological constituency of maximal syllable margin: Morpheme-internal, Morphologically Complex, or Both patterns occur

Morphological pattern of syllabic consonants: Lexical items, Grammatical items, Both

Additionally, more detailed information on syllable patterns, restrictions on cluster patterns, syllable type frequency data, and the phonetic characteristics of clusters was gathered for the Highly Complex portion of the sample.

An example of the syllable structure coding for Ket, a language with Complex syllable structure, can be found below (11).

(11) **Ket** (*Yeniseian*; Russia)

Size of maximal onset: 2

Size of maximal coda: 3

Onset obligatory: No

Coda obligatory: No

Vocalic nucleus patterns: Short vowels

Syllabic consonant patterns: Nasals

Size of maximal word-marginal sequences with syllabic obstruents:

N/A

Predictability of syllabic consonants: Predictable from word/consonantal context

Morphological constituency of maximal syllable margin:

Morphologically Complex (Onsets), Both patterns (Codas)

Morphological pattern of syllabic consonants: Grammatical items

All syllable structure coding for the language sample, including illustrative examples and notes on specific onset and coda patterns, can be found in Appendix B.

1.3 Results for language sample

In this section I present analyses of general patterns of maximal onsets, codas, and syllabic consonants in the language sample. In §1.3.1 I present the maximal onset and coda sizes in the data, and discuss maximal word-marginal patterns in

the languages which have syllabic obstruents. In §1.3.2, I briefly examine the relationship between onset and coda complexity. In §1.3.3 the relationship between syllable structure complexity and obligatoriness of syllable margins is investigated. In §1.3.4 and §1.3.5 the patterns of vocalic nuclei and syllabic consonants, respectively, are presented and analyzed with respect to syllable structure complexity. §1.3.6 is a longer subsection which addresses the issue of morphological constituency in maximal syllable margins and syllabic consonants in the data.

1.3.1 Maximal onset and coda sizes

The distribution of maximal onset and coda patterns in the sample, by number of consonants, can be found in Figures 1.1–1.2.

Figure 1.1: Maximal onset sizes in sample.

Figure 1.2: Maximal coda sizes in sample.

The data presented in Figures 1.1 and 1.2 are for onset and coda patterns determined through the procedure described in §1.2.1. They do not include the largest word-marginal patterns which occur in languages with syllabic obstruents. There are four languages in the sample which are reported to have syllabic obstruents resulting in Highly Complex patterns at word margins.⁴ These patterns are presented in Table 1.1, along with the reported maximal onset and coda patterns in the languages.

In Tashlhiyt, the size of maximal word-marginal obstruent strings cannot be determined, because there are many examples of words consisting entirely of obstruents in this language (12).

- (12) **Tashlhiyt** (*Afro-Asiatic*; Morocco)
tftktst:
tf.tk.tst:
 ‘you took it off (F)’ (Ridouane 2008 : 332)

In Tehuelche, the reference indicates that sequences of up to six consonants may occur word-finally, but the only illustrative example of a pattern this size is given is (13).

⁴Additionally, Tohono O’odham has syllabic obstruents, but only in independent grammatical particles consisting of a single consonant (determiners and conjunctives). These are not reported to be phonologically bound to adjacent words. Therefore I do not include Tohono O’odham in Table 1.1.

1 Syllable structure patterns in sample

Table 1.1: Languages in Highly Complex category with syllabic obstruents. Maximal reported onsets and codas are given in first two columns. The sizes of the maximal word-marginal obstruent strings which occur as the result of syllabic obstruents are given in last two columns.

Language	Maximal onset	Maximal coda	Maximal obstruent string	
			word-initial	word-final
Cocopa	4	3	5	3
Semai	2	1	4	1
Tashlhiyt	1	1	(words without vowels)	
Tehuelche	2	3	3	>3

(13) **Tehuelche** (*Chonan*; Argentina)

kʰʃaʔʃpʰk'n

k.ʃaʔʃp.f.k'n

‘it is being washed’

(FernándezGarayHernández2006: 13)

This example includes a nasal which may be syllabic (syllable peaks in CC syllables are not marked by the authors). It is clear from the language description that long obstruent sequences come about when syllabic consonants are strung together, but unclear as to what the upper limit on the size of these is. Word-final sequences of at least four obstruents are attested (14).

(14) **Tehuelche** (*Chonan*; Argentina)

ma:leʃpʰk'

ma:.leʃp.f.k'

‘they steal’

(FernándezGarayHernández2006: 63)

1.3.2 Relationship between onset and coda complexity

Here I present an analysis similar to the one presented in §?? for the Complex portion of the Maddieson (2013) sample. The languages of the sample used in this book are distributed according to their maximal onset and coda patterns.

Interestingly, it is common for languages with large clusters at one syllable margin to also exhibit large clusters at the other syllable margin in their canonical patterns. Roughly half of languages in the sample which have a maximal

Table 1.2: Languages of sample distributed according to maximal onset and coda size.

Number of Cs in coda	Number of Cs in onset							
	One	Two	Three	Four	Five	Six	Seven	Eight
None	20	1	2	1	—	—	—	—
One	21	12	5	1	—	—	—	—
Two	7	5	5	—	—	—	—	—
Three	1	4	2	4	—	—	—	—
Four	3	—	—	2	—	—	—	—
Five	—	—	—	—	1	—	1	1
Six	—	—	1	—	—	—	—	—

cluster of four or more consonants at one syllable margin will have a similarly large maximal cluster at the other syllable margin. It is striking that *all* languages in the sample with maximal onsets of five or more consonants (Georgian, Itelmen, and Polish) also have maximal codas of five consonants. Meanwhile, the bottom left and top right corners of Table 1.2 are sparsely populated; that is, there are relatively few languages with very large maximal clusters at one syllable margin and very small maximal clusters (or none at all) at the other margin. A similar pattern can be observed in Table ?? in §??, which used a larger sample of 147 languages. Speaking from a strictly distributional point of view, there is no obvious motivation for this pattern. If we consider onset and coda structures to be independent structures, then we would expect to see the full range of possible variation in their combination crosslinguistically. This point will be revisited in §1.5 and again in Chapter ??.

1.3.3 Syllable structure complexity and obligatoriness of syllable margins

Blevins (2006 : 336) notes another crosslinguistic pattern linking onset and coda structures: she observes that languages with only open syllables tend to have optional onsets. In Table 1.3 I examine this relationship in the current language sample. There are some languages in which optional onsets may be reported for the canonical syllable structure, but regular and obligatory consonant epenthesis (usually of a glottal stop or fricative) occurs to produce onsets in all ‘surface’

1 Syllable structure patterns in sample

forms. In the analysis below, I consider such languages as having obligatory onsets.

Table 1.3: Languages in sample distributed according to occurrence of codas and obligatoriness of onsets.

<i>N</i> languages	Codas occur?	Onset obligatory?
14	Y	Y
62	Y	N
2	N	Y
22	N	N

The relationship reported by Blevins is upheld in the current sample. Of the 24 languages with only open syllables (no codas), only two are reported to have obligatory onsets. Obligatory onsets are more common in languages with coda structures (14/76). The two languages in the sample with obligatory onsets and only open syllables are Hadza (Simple) and Yine (Highly Complex). It should also be noted that the Oykangand dialect of Kunjen used here is reported to have obligatory codas. Kunjen is argued to have very marginal onset patterns, with onsets occurring in interjections and sentence-initially in a just a few lexical items (Sommer 1969 ; 1970 ; 1981, though see Dixon 1970 for an opposing view).⁵ Therefore Kunjen shows a very similar pattern to Hadza and Yine, except that the syllable margins are reversed.

The analysis above motivated a more general examination of obligatory syllable margin patterns in the language sample with respect to syllable structure complexity (Table 1.4).

Obligatory syllable margins are a minor pattern in the language sample, occurring in only 17 languages, but this feature is most common in languages with Highly Complex syllable structure, occurring in roughly one-third (8/25) of those languages. This feature is least common in languages with Simple syllable structure (2/24 languages). The pattern in the Highly Complex category is statistically significant when compared against the patterns in the other three categories combined ($p = .03$ in Fisher's exact test). This association between syllable structure complexity and obligatoriness of syllable margins has, to my knowledge, not

⁵According to Sommer's analysis, Kunjen is a rare example of a language without phonological CV syllables. Another language argued not to have phonological CV syllables is Arrernte, a Pama-Nyungan language of central Australia (cf. Breen & Pensalfini 1999), though Anderson (2000) reports a canonical surface syllable structure of (C)(C)V(C) for Western Arrernte.

Table 1.4: Languages in sample with obligatory syllable margins.

	Syllable structure complexity			
	Simple N = 24	Moderately Complex N = 26	Complex N = 25	Highly Complex N = 25
Onset obligatory	<i>Hadza</i> <i>Ute</i> (2)	<i>Kambaata</i> <i>Karok</i> <i>Lao</i> <i>Pacoh</i> (4)	<i>Koho</i> <i>Lepcha</i> <i>Mangarrayi</i> (3)	<i>Bench</i> <i>Nuu-chah-nulth</i> <i>Semai</i> <i>Thompson</i> <i>Tohono O'odham</i> <i>Yakima Sahaptin</i> <i>Yine</i> (7)
Coda obligatory	—	—	—	<i>Kunjen</i> (1)

been previously reported.

Obligatory syllable margins are more common in some areas than others: Five of the languages in Table 1.4 are from Southeast Asia & Oceania. North America is also heavily represented, accounting for another six languages altogether, including four of the languages with obligatory syllable margins from the Highly Complex group. It should further be noted that most of the North American languages with obligatory syllable margins in the Highly Complex group are from the Pacific Northwest (Nuu-chah-nulth, Thompson, and Yakima Sahaptin), so areal factors may be at play. Nevertheless, the Highly Complex pattern in Table 1.4 is not entirely areal in nature, as it includes languages from Africa, Southeast Asia & Oceania, South America, and Australia & New Guinea.

Interestingly, the description of Itelmen (Chukotko-Kamchatkan, Highly Complex) suggests that this language, too, had obligatory onsets at some point in its history: morphophonological processes suggest that present-day vowel-initial syllables were at one time initiated by a glottal stop (Georg & Volodin 1999 : 48).

1.3.4 Vocalic nucleus patterns

Vocalic nucleus patterns have until now been excluded from the discussion of syllable patterns, as they are not considered in the definitions of syllable structure complexity used here. However, it is important to note that vocalic nucleus

1 Syllable structure patterns in sample

patterns can also exhibit different degrees of complexity. In Table 1.5 I present a very general analysis of these patterns in the sample, showing the distribution of simple and complex vocalic nuclei by syllable structure complexity.

Table 1.5: Vocalic nucleus patterns in language sample, by syllable structure complexity.

Languages with:	Syllable Structure Complexity			
	Simple N = 24	Moderately Complex N = 26	Complex N = 25	Highly Complex N = 25
Simple vocalic nuclei only	12	9	9	9
Complex vocalic nuclei ^a	12	17	16	16

^a(long vowels, diphthongs, and/or vowel sequences)

Simple vocalic nuclei – those consisting of a single short vowel – occur in every language. The first row of Table 1.5 shows the number of languages in each complexity category for which this is the only vocalic nucleus pattern occurring. Languages in which complex vocalic nuclei occur in addition to simple vocalic nuclei are shown in the second row. For the sake of simplicity I have collapsed three different kind of complex vocalic nucleus patterns in the analysis here. A language is counted as having long vowels if it has contrastive vowel length, but not if it has predictable vowel lengthening, e.g., a longer variant preceding a voiced coda. Diphthongs and tautosyllabic vowel sequences are difficult to disambiguate from one another, as their analyses by different authors may vary widely; however, vowel sequences reported here as syllable nuclei are those explicitly shown by the author to belong to one syllable, much like a diphthong. That is, this figure does not include cases of hiatus, in which the two vowels in a sequence belong to different syllables.

Table 1.5 shows that complex vocalic nuclei are much less likely to occur in languages with Simple syllable structure than in languages from the other categories. This suggests that the potential for more syllable types in languages with more complex syllable structure may be not only a function of larger canonical syllable margins, but also of greater diversity in syllable nucleus patterns. Nevertheless, the analysis above is too coarse to draw strong conclusions about vocalic nucleus patterns and syllable structure complexity. The issue of contrastive vowel length will be treated in greater detail in Chapter ??, along with contrastive nasalization, voicing, and glottalization patterns in the vowel inventories of the sample.

1.3.5 Syllabic consonants

In this section I investigate patterns of syllabic consonants in the data. Recall that the previous literature suggests two competing predictions for the relationship between syllable complexity and the presence of syllabic consonants. Isačenko's (1939) phonological typology predicts that 'vocalic' languages, which tend to have simpler syllable structure, will be more likely to develop syllabic consonants, and specifically syllabic sonorants. Meanwhile, Bell (1978) notes that syllabic consonants, including syllabic obstruents, often come about through vowel reduction processes, which are also known to produce the clusters characteristic of languages with more complex syllable structure. On the basis of the latter observation, in §1.1.2 I formulated a hypothesis that languages with more complex syllable structure are more likely to have syllabic consonant patterns.

Here I analyze languages in which the syllabic consonants are reported as invariant patterns. Most often, the syllabicity of these consonants is predictable from the surrounding consonantal and/or word environment, as illustrated by (15). Less frequently, syllabic consonants are analyzed as separate phonemes which are contrastive with their non-syllabic counterparts (16a-b).

- (15) **Itelmen** (*Chukotko-Kamchatkan*; Russia)

A word-initial alveolar or bilabial nasal stop preceding another consonant is realized as syllabic.

/m̥im/

[m̥im]

'blood'

(Georg & Volodin 1999 : 16)

- (16) **Ewe** (*Atlantic-Congo*; Ghana, Togo)

a. jɔm̥

'call me'

b. kampe'

'scissors' (Ameka 1991 : 38)

check
diacritics

Three languages are excluded from the current analysis: Chipaya, Nimboran (both from the Complex category), and Yine (from the Highly Complex category). For all three of these languages, there are conflicting reports regarding the oc-

1 Syllable structure patterns in sample

currence of syllabic consonants, sometimes from the same author.⁶ For example, Matteson gives the following description for Yine, which seems to suggest that consonants in complex onsets both belong to a syllable with a vocalic nucleus and are simultaneously themselves syllabic:

“We number the consonants of the syllable, beginning with the consonant that immediately precedes the nuclear vowel: +C³ +C² +C¹V. In the positions of consonants C² and C³ occur syllabic allophones of the consonants. Thus the syllable is a complex unit consisting of from one to three syllabic units.” (Matteson 1965 : 23)

Because of the conflicting descriptions of these languages, I opted to exclude them from the current analysis. Georgian and Tashlhiyt also have conflicting descriptions with respect to the occurrence of syllabic consonants, but in both of these cases experimental evidence has been presented to support one analysis over another. The articulatory and acoustic experiments in Ridouane (2008) and Goldstein et al. (2007) support a syllabic consonant analysis for Tashlhiyt, while native speaker intuition reported in Chitoran (1999) does not support an analysis of syllabic sonorants for Georgian. It is interesting to note that all of the languages with conflicting descriptions – those discussed here and the ones mentioned in the footnote – are from the Complex and Highly Complex categories. This recalls the observation noted previously, in which transitions in consonant clusters on the one hand and syllabic consonants on the other may have similar motivations and acoustic manifestations.

The syllabic consonant patterns reported for the languages of the sample can be found in Table 1.6. There is a steadily increasing trend in the proportion of languages with these patterns as syllable structure complexity increases: only two of languages with Simple syllable structure are reported to have syllabic consonants, compared to 11 of the languages in the Highly Complex category. The

⁶There are a few other languages for which there are suggestions of alternate analyses. The dialect of Sahaptin analyzed here, Yakima, is argued not to have syllabic consonants by Hargus & Beavert (2006) on the basis of distributional and phonological behavior of consonants in sequences. However, it should be noted that Minthorn (2005) argues for syllabic consonants, including obstruents, in the closely related dialect of Umatilla Sahaptin, on the basis of speaker intuition and acoustic analysis. Additionally, one description of Alamblak lists an example of a word consisting entirely of obstruents: *kpt* ‘basket type’ SIL2004; however, no further elaboration is given and obstruents are not included in the description of syllabic consonants in Bruce (1984), so it is unclear whether syllabic obstruents are an issue of debate for this language. Finally, for Itelmen, Volodin (1976 : 42) gives transcriptions of lexical items consisting entirely of obstruents (*ʃʃkpfʃ* ‘spoon’). In a later reference, he describes only syllabic sonorants in the language (Georg & Volodin 1999).

Table 1.6: Presence of invariant syllabic consonants in language sample, by syllable structure complexity. Chipaya, Nimboran (Complex) and Yine (Highly Complex) excluded.

Languages with:	Syllable Structure Complexity			
	Simple N = 24	Moderately Complex N = 26	Complex N = 23	Highly Complex N = 24
Syllabic consonants				
of any kind	2	6	5	11
<i>Syllabic nasals</i>	2	5	5	10
<i>Syllabic liquids</i>	—	2	1	6
<i>Syllabic obstruents</i>	—	—	—	5
No syllabic consonants	22	20	18	13

trend in the Highly Complex category is statistically significant when compared to the trends in the other three categories combined ($p = .01$ in Fisher's exact test). Examining the particular kinds of syllabic consonants represented, the patterns are similar to what is reported in [Bell \(1978\)](#). Most languages with syllabic consonants have syllabic nasals, and languages with syllabic obstruents are rare. While languages from all four categories have syllabic nasals and most have syllabic liquids, syllabic obstruents are only reported for languages in the Highly Complex category. This is not a remnant of the way the Highly Complex category is defined: recall that languages with syllabic obstruents are categorized as Highly Complex only if these structures participate in word-marginal sequences of three obstruents or more. It is striking that no languages with simpler syllable structure are reported to have syllabic obstruents. Even if the three languages excluded from the previous analysis were included here, the distribution of syllabic obstruents would be among two languages with Complex syllable structure and six languages with Highly Complex syllable structure.

It should also be noted that three of the languages with syllabic obstruents (Cocopa, Semai, Tashlhiyt) are reported to also have both syllabic nasals and syllabic liquids. Tehuelche does not have syllabic liquids. Tohono O'odham is the only language which has syllabic obstruents but not syllabic sonorant consonants. This indicates that the trend in [Table 1.6](#) — by which Highly Complex languages are more likely than any of the other categories to have syllabic consonants — is not driven by the inclusion of syllabic obstruents in the definition of that category, or

1 Syllable structure patterns in sample

skewed by potential misanalyses which confound syllabic obstruents and large tautosyllabic clusters. Instead, the trend can be obtained from the syllabic nasal and liquid patterns in the sample.

While the analyses presented above are for the invariant syllabic consonant patterns observed in the sample, there were also several cases in which syllabic consonants were reported to occur in variation with CV or VC sequences, as illustrated by (17)–(18).

- (17) **Sichuan** (; Yi)
 Sino-TibetanChina
Nasals and laterals preceding [i] occur in free variation with syllabic consonants.
 /li/
 [li]~[ɭ] (Gerner 2013 : 31)
- (18) **Mamaindê** (Nambiquaran, Brazil)
When an unstressed vowel is lost resulting in a sequence of nasal plus consonant, a preceding nasal becomes syllabic.
 /ˈjohnalat^hawa/
 [ˈjohnɫat^hwa]
 ‘it is low’ (Eberhard 2009 : 262-3)

Table 1.7: Distribution of languages in sample with syllabic nasals, liquids, and obstruents occurring in variation with VC or CV structures.

Languages with variable:	Syllable Structure Complexity			
	Simple N = 1	Moderately Complex N = 3	Complex N = 2	Highly Complex N = 3
<i>Syllabic nasals</i>	1	3	2	3
<i>Syllabic liquids</i>	1	—	—	2
<i>Syllabic obstruents</i>	—	1	—	1

Table 1.7 shows the distribution of variable processes producing syllabic consonants in the data. Though the data set is very small, it is interesting that the general distributional pattern is similar to that presented in Table 1.6. The occurrence of syllabic consonants in variation with VC or CV structures is least frequent among languages with Simple syllable structure, and in all categories

nasals are the most common syllabic consonant to result. Variable syllabic obstruents occur in two languages. In Paiwan (Moderately Complex), syllabic obstruents may occur when schwa is reduced immediately after a sibilant in rapid speech, and in Kabardian (Highly Complex), they occur as the result of an optional process of high vowel contraction (3.19-3.20):

- (19) **Paiwan** (*Austronesian*; Taiwan)

/səkam/

[səkam]~[ʃkam]

‘mattress’ (Chang 2006 : 41)

- (20) **Kabardian** (*Abkhaz-Adyge*; Russia, Turkey)

/tʰəʒ/

[tʰiʒ]~[tʰʒ]

‘old man’ (Kuipers 1960 : 24)

Kabardian is also the only language in the sample reported to have both invariant syllabic consonants (for sonorants in certain consonant environments) and variable syllabic consonants as a result of synchronic phonetic processes like the one illustrated above.

Returning to the hypothesis stated at the beginning of this section, there is evidence that languages with more complex syllable structure are more likely to have syllabic consonant patterns. Specifically, languages with Highly Complex syllable structure are the most likely of all those in the sample to have invariant syllabic consonants, while languages with Simple syllable structure are the least likely. Variable processes resulting in syllabic consonants are also relatively more frequent in languages with non-Simple syllable structure.

1.3.6 Morphological patterns

In this section I analyze the morphological patterns associated with syllable patterns in the language sample. First, I report the morphological constituency patterns observed in the maximal onset and coda structures in each language. Then I present an analysis of the kinds of morphemes (lexical or grammatical) in which syllabic consonants in the language sample occur. I test the hypotheses formulated in §1.1.2 with respect to these patterns: first, that as syllable structure complexity increases, so does the likelihood that the largest syllable margin types in a language will be morphologically complex; second, that as syllable structure

1 Syllable structure patterns in sample

complexity increases, so does the likelihood that syllabic consonants occurring in a language will belong to grammatical elements.

Since morphologically complex instances of syllable patterns are often not explicitly described and must be gathered from the examples, it was impractical and in many cases impossible to find morpheme-internal and morphologically complex instances of the same specific consonant sequence in each language, especially for the larger clusters. The patterns analyzed here are for the maximal onset and coda *types*, e.g. CC. For example, the maximal coda in Gaam is two consonants. The word-final patterns shown in the examples below would be taken as evidence that the maximal coda occurs in both morpheme-internal (21a) and morphologically complex (21b) contexts.

check
diacritics

(21) **Gaam** (*Eastern*;)
JebelSudan

a. bã gɗa`rs

‘lizard type’

b. gəũ r-ɗ

stomach-sg

‘stomach’ (Stirtz 2011 : 32, 37)

Note also that the definition of *morphologically complex* here refers to sequences derived by any morphological process. That is, sequences derived through reduplication or nonconcatenative processes such as subtractive morphology are also considered to be morphologically complex, even though they don’t involve more than one distinct morpheme.

First I test **Greenberg’s** (1965) prediction in the data: as the size of a syllable margin increases, so does the probability that it contains morpheme boundaries. Figures 1.3 and 1.4 show morphological constituency patterns in maximal onset and coda types in the data.

Figure 1.3: Morphological constituency patterns in maximal onset types in data. For each maximal onset type, figure shows proportion of languages exhibiting the given morphological patterns for that type.

Figure 1.4: Morphological constituency patterns in maximal coda types in data. For each maximal coda type, figure shows proportion of languages exhibiting the given morphological patterns for that type.

For both maximal onset and maximal coda patterns in the data, the proportion of languages having these clusters solely in morphologically complex contexts

increases with cluster size. However, morphologically complex patterns also occur alongside morpheme-internal patterns in the maximal margins for a number of languages (the “Both patterns” trend in Figures 1.3 and 1.4). When this trend is additionally considered, we find that maximal coda cluster types are generally more likely than maximal onset types to exhibit morphologically complex patterns. We also find that all maximal cluster types of five consonants or larger are found only in morphologically complex contexts.

Interestingly, there are some language-internal patterns in the data which go against Greenberg’s prediction. In Lelepa there are biconsonantal onsets showing both morphological patterns, but the only attested triconsonantal onsets are within morphemes (22a-c).

(22) **Lelepa** (*Austronesian*; Vanuatu)

- a. n-malogo
NMLZ-darken
‘darkness’
- b. nmal
‘trunk’
- c. psruki
‘speak’

(Lacrampe 2014 : 107, 207, 42)

The analyses presented in Figures 1.3–1.4 test Greenberg’s specific predictions regarding cluster size. However, the hypothesis in (4a) is formulated with respect to syllable structure complexity, which is a slightly different question, though we expect to find a similar pattern due to how the categories are defined. In Figures 1.5–1.6 I present the morphological constituency patterns observed by syllable structure complexity category. Note that these figures only include the languages from each category which have complex onsets or complex codas, respectively.

??

Figure 1.5: Morphological constituency patterns in maximal complex onsets, by syllable structure complexity category. For each category, the figure shows the proportion of languages exhibiting the given morphological patterns in its complex onsets.

The figures show that as syllable structure complexity increases, both maximal onset and maximal coda clusters are more likely to have morphologically complex patterns, confirming the hypothesis in (4a).

The patterns in figures 3.5–3.6 are combined in Figure 1.7 in order to show the general trend for morphologically complex patterns in maximal syllable-marginal

1 Syllable structure patterns in sample

Figure 1.6: Morphological constituency patterns in maximal complex codas, by syllable structure complexity category. For each category, the figure shows the proportion of languages exhibiting the given morphological patterns in its complex onsets. Note that the one language with complex codas from the Moderately Complex category — Eastern Khanty — is not included in the figure. Its very marginal complex codas are always morphologically complex.

clusters with respect to syllable structure complexity in the language sample. In this figure onset and coda patterns are collapsed, and the “both contexts” and “only morphologically complex” patterns are combined. For each category, the figure shows the percentage of languages with complex syllable margins for which morphologically complex patterns occur in either or both maximal syllable margins.

Figure 1.7: Percentage of languages in each category exhibiting morphologically complex patterns in either or both of its maximal syllable margins.

Morphologically complex patterns can be found in the maximal syllable margins of most languages from the Highly Complex category, and this pattern is statistically significant when compared against the patterns in the other two categories combined ($p < .01$ in Fisher’s exact test). However, there are six languages in this category for which I could determine only morpheme-internal patterns. In Wutung, the maximal margin is explicitly described as occurring within a few apparently single-morpheme lexical items. In Kunjen and Lezgian, maximal coda and onset clusters, respectively, seem to be limited to single-morpheme lexical items, though the references consulted do not explicitly state this. In Menya, the only morphologically complex instance of a maximal cluster that could be found was in an abstract phonemic transcription for which the phonetic form was unclear. In Passamaquoddy-Maliseet, examples of morphologically complex instances of maximal clusters could not be found, though it seems as though the morphology could produce them. The remaining language, Semai, has syllabic consonants and will be discussed below.

Recall that there are four languages in the Highly Complex portion of the sample for which the largest word-marginal obstruent sequences include syllabic consonants. The maximal ‘true’ onset/coda clusters reported for these languages (cf. Table 1.2) were included in the previous analyses in this section, but the maximal word-marginal sequences were not. I present the morphological patterns for these sequences in Table 1.8.

All of the maximal word-marginal obstruent sequences in the languages in

Table 1.8: Morphological patterns of maximal word-marginal obstruent sequences in languages with syllabic obstruents in Highly Complex category.

Language	Maximal word-initial obstruent string	Morphological pattern	Maximal word-final obstruent string	Morphological pattern
Cocopa	5	complex	3	complex
Semai	4	complex	1	—
Tashlhiyt	— ^a	complex	— ^a	complex
Tehuelche	3	complex	>3	complex

^a(words without vowels)

Table 1.8 occur in morphologically complex contexts. For example, in Semai, all maximal word-initial consonant sequences, and indeed all word-initial sequences of more than two consonants, are derived through reduplication processes (23).

- (23) **Semai** (*Austroasiatic*; Malaysia)
gp.g.hup (< ghup)
‘irritation on skin (e.g., from bamboo hair)’
(Sloan 1988 : 320; Diffloth 1976 : 256)

Though maximal word-marginal obstruent string length cannot be determined in Tashlhiyt due to the occurrence of many words consisting entirely of obstruents in this language, the longest such words are morphologically complex (24).

- (24) **Tashlhiyt** (*Afro-Asiatic*; Morocco)
t-s:-kff-t=st:
ts.sk.ff.tst:
‘you dried it (F)’
(Ridouane 2008 : 332; interlinear gloss not provided)

Now we turn to the hypothesis in (4b): as syllable structure complexity increases, so does the likelihood that syllabic consonants occurring in a language will belong to grammatical elements. This is based partly on Bell’s (1978) observation that the syllabic consonants in his typological survey were often restricted to grammatical particles and affixes.

Only the languages reported in §1.3.5 as having invariant syllabic consonant patterns are included in the analysis here. Additionally, Kabardian is excluded

1 Syllable structure patterns in sample

from the present analysis because its precise patterns could not be determined. For each kind of syllabic consonant analyzed (nasal, liquid, and obstruent), I determine whether that type occurs in lexical morphemes, grammatical morphemes, or both. For example, in Bench, syllabic nasals can be found in both lexical and grammatical morphemes (25).

(25) **Bench** (*Ta-Ne-Omoti*c; Ethiopia)

- a. gũpm̐
‘foam’
- b. njãʔ-n̐ d
child-PL
‘children’
(Rapold 2006 : 107, 111)

In Tables 1.9–1.11 I present analyses for the morphological patterns of each kind of syllabic consonant (nasal, liquid, and obstruent) observed in the data.

Table 1.9: Morphological patterns of syllabic nasals in sample, by syllable structure complexity. Kabardian (Highly Complex) is omitted as its pattern could not be determined.

Languages with	Syllable structure complexity			
	Simple	Moderately Complex	Complex	Highly Complex
syllabic nasals in:	<i>N</i> = 2	<i>N</i> = 5	<i>N</i> = 6	<i>N</i> = 9
<i>Lexical morphemes only</i>	1	3	2	1
<i>Lex. and gram. morphemes</i>	—	2	2	6
<i>Grammatical morphemes only</i>	1	—	2	2

In general, the pattern by which syllabic consonants are found to occur in grammatical morphemes, either exclusively or in addition to lexical morphemes, is the dominant one in the data. This is the case for 15/22 languages with syllabic nasals and all of the languages with syllabic obstruents. And within this very small data set, this trend also appears to increase with syllable structure complexity, suggesting support for the hypothesis.

In fact most of the languages with syllabic consonants in the Highly Complex category have these sounds in grammatical items. In Tehuelche, for example, *all* syllabic consonants correspond to or belong to grammatical morphemes (26).

1.3 Results for language sample

Table 1.10: Morphological patterns of syllabic liquids in sample, by syllable structure complexity. Kabardian (Highly Complex) is omitted as its pattern could not be determined.

Languages with syllabic liquids in:	Syllable structure complexity			
	Simple <i>N</i> = 0	Moderately Complex <i>N</i> = 2	Complex <i>N</i> = 1	Highly Complex <i>N</i> = 5
<i>Lexical morphemes only</i>	—	2	1	2
<i>Lex. and gram. morphemes</i>	—	—	—	2
<i>Grammatical morphemes only</i>	—	—	—	1

Table 1.11: Morphological patterns of syllabic obstruents in sample, by syllable structure complexity.

Languages with syllabic obstruents in:	Syllable structure complexity			
	Simple <i>N</i> = 0	Moderately Complex <i>N</i> = 0	Complex <i>N</i> = 0	Highly Complex <i>N</i> = 5
<i>Lexical morphemes only</i>	—	—	—	—
<i>Lex. and gram. morphemes</i>	—	—	—	1
<i>Grammatical morphemes only</i>	—	—	—	4

- (26) **Tehuelche** (*Chonan*; Argentina)
 k.ʃjaʔʃp.ʃ.k'n
 k-ʃjaʔʃp-f-k'n
 REFL-wash-PS-REALIS
 'it is being washed'
 (FernándezGarayHernández2006: 13)

To summarize, in this section the morphological patterns of maximal onset types, maximal coda types, and syllabic consonant inventories in the sample have been examined. In both cases there is support for the hypotheses in (4). Clearly morphology contributes an important role to the development of complex syllable patterns. While this point will be only briefly revisited in the discussion in §1.5, it will be discussed in further detail in Chapter ??.

1.4 Properties of highly complex syllable structure

Having described the general patterns of maximal onsets, maximal codas, and syllabic consonants in the data, I now turn to an examination of the properties of syllable structure in the languages in the Highly Complex portion of the sample. In §1.4.1 I give examples of the syllable patterns occurring in each of these languages. In §1.4.2 I attempt to characterize the prevalence of Highly Complex structures within each of the languages by examining restrictions on consonant combinations and reported frequency patterns. In §1.4.3 I present information on the acoustic and perceptual properties of Highly Complex structures.

1.4.1 Examples of Highly Complex syllable patterns in sample

In order to provide a better picture of what specific syllable patterns occur in the languages of the Highly Complex portion of the sample, I list some representative structures in Table 1.12. The definition of Highly Complex syllable structure includes any onset or coda structure of three obstruents, or of four consonants or more in length. It also includes any word-marginal sequence containing syllabic obstruents such that a sequence of three or more obstruents occurs at a word margin. For each language I give a set of examples for each onset, coda, and/or word-marginal cluster that occurs at each of the following lengths: three consonants, four consonants, and five or more consonants. For Tashlhiyt, I have given some examples of vowelless words in the rightmost column, but have not assigned them to a word margin.

The languages in Table 1.12 are organized so as to highlight several coherent patterns in the data. In the first set of languages (Alamblak, Bench, Menya, Kabardian, Lezgian, Yine, Camsá, Semai, and Nuu-chah-nulth), Highly Complex patterns are limited to one syllable or word margin, usually the onset/initial context. The Highly Complex patterns in these languages are typically limited to triconsonantal clusters, though four-consonant clusters occur in Camsá, Semai, and Nuu-chah-nulth. In the second group of languages (Wutung, Doyayo, and Kunjen), four-consonant clusters occur at one syllable margin, but triconsonantal patterns falling under the definition of Highly Complex (that is, sequences of three obstruents) do not occur. In this group, the CCCC clusters include at least one, but usually two, sonorants. Finally, in the remaining 13 languages, Highly Complex patterns occur in both margins and almost always include clusters of various sizes.

It is typically the case in the language sample that if a language has syllable margins of three obstruents, then any larger margins which occur in the lan-

1.4 Properties of highly complex syllable structure

Table 1.12: Representative sample of Highly Complex patterns occurring in data. (—) indicates that there are no reported patterns of this kind in the given language.

Language	Highly Complex structures		
	3-obstruent	4-C	5-C and larger
Alamblak	Onset: tkb	—	—
Bench	Coda: pst	—	—
Menya	Onset: tpq, ptq	—	—
Kabardian	Onset: zby, pət, psk'	—	—
Lezgian	Onset: ftk, kst, ktk	—	—
Yine	Onset: (pcr, nŋsp, nŋʃk)	—	—
Camsá	Onset: stx, sŋʃb, sŋʃs Onset: φstx	—	—
Semai	Word-initial: st.s	Word-initial: gp.g.h	—
Nuu-chah-nulth	Coda: ʃʃtq, kqs, qʃʃs, thʃs	Coda: mtqʃ, hsqh, nkqh	—
Wutung	—	Onset: hmb	—
Doyayo	—	Coda: βlts, yldz, mnts	—
Kunjen	—	Coda: lbmb, ɾnd, jɔŋg	—
Passamaquoddy-Maliseet	Onset: psk, ksp, psk ^w Coda: psk ^w , ksk ^w	—	—
Qawasqar	Onset: qsq, qst, qsk Coda: qsq	Onset: qsqj	—
Tehuelche	Word-initial: kʃ.x, kʃ.ʔ Coda: ʔʃp	Word-final: ʃp.ʃ.k'	—
Albanian	Onset: skt, pʃt Coda: pʃt, kst	Onset: ʃmpl, zmbr	—
Mohawk	Onset: ksk, kts, kst, kht Coda: ʔks, ʔts, kst	Onset: shnj khnj	—
Yakima Sahaptin	Onset: pʃχ, tk ^w s, q'ʃp Coda: tks, stk, pʃʃ'k	Onset: ʃtχn, ksks Coda: wtk ^w ʃ, wq'χʃ, jlp	—
Tohono O'odham	Coda: gsp, tpk, bʃʃ, psk	Onset: ndʃʔ Coda: ʃʃʃktʃ, ʃʃspk, gssp	—
Polish	Onset: pʃt, xʃt, tkʃ Coda: psk, stf, ʃʃp	Onset: pstʃ, fksʃ, vzgl Coda: ɲstf, tstf, rstf, pstf	Onset: spstr Coda: mpstr

The Yine patterns are in parentheses because they are representative triconsonantal clusters for the language but do not contain three obstruents (see discussion in §1.2.3).

1 Syllable structure patterns in sample

Table 1.13: Representative sample of Highly Complex patterns occurring in data. (—) indicates that there are no reported patterns of this kind in the given language. The Yine patterns are in parentheses because they are representative triconsonantal clusters for the language but do not contain three obstruents (see discussion in §1.2.3).

Language	Highly Complex structures		
	3-obstruent	4-C	5-C and larger
Thompson	Onset: spt, sfs'k Coda: x ^w kt, x ^w sfs, psfs	Coda: fsxsfs, jxsfs, fkst	Coda: fqsxtx ^w
Itelmen	Onset: kth kp'k' fqtz Coda: plh sht	Onset: ttxn, ksxw, ktxl Coda: ntfx, mptx, ftxf	Onset: kptkn, tksxqz, k Coda: nxftx, mstxtf
Georgian	Onset: t'k'b p'fs'k' psk'	Onset: txzβ, fs'q'rt, brfs'q' Coda: rtxl, rt'q'l, ntfxl	Onset: p'rts'k'β, gβ, p Coda: ntfxls, rfs'q'β, s, f
Cocopa	Onset: sxt, psk ^w , xps Coda: qsk, fsk, xsk	Onset: ffx? pff?, p.fxm	Word-initial: pk.fkw
Tashlhiyt	Word-initial: ts.t Word-final: k ^w tt, f.kd	Word-initial: ts:χs Word-final: stst:	(Words without vowels)

guage may also include sequences of three or more obstruents. The only apparent exceptions to this trend are four-consonant onsets in Albanian and Mohawk, and four-consonant codas in Georgian. In these cases, the larger clusters always include more than one sonorant, such that sequences of more than two obstruents do not occur, e.g. Georgian coda /rt'q'l/. In all other languages with both triconsonantal and larger Highly Complex structures, long strings of obstruents are a hallmark characteristic of the larger structures. That is, the patterns in the third group of languages described above are not simply an amalgamation of the patterns from the first and second groups of languages described above. The second group (Wutung, Doyoyo, and Kunjen) represents a minority pattern in that the only Highly Complex structures occurring in these languages do not involve strings of more than two obstruents.

It should also be noted that languages with syllabic consonants do not behave as a group apart from the other languages with respect to the distribution of their Highly Complex sequences. Semai patterns with the first group of languages, while Tehuelche, Cocopa, and Tashlhiyt pattern with the third group.

Table 1.12 does not provide an exhaustive list of Highly Complex structures for each language; however, for a few languages for which this is a minor pattern,

an exhaustive or near-exhaustive list is given. This is the case for Alamlak and Menya. The Highly Complex onsets listed for these languages are not explicitly stated in the references to be the only structures of this sort, but a search of the examples and texts yielded only these patterns. In Bench and Wutung, the single onset given for each language is explicitly stated to be the only one occurring. For other languages, the lists given for larger structures may be exhaustive, but those given for smaller structures may be a tiny representative sample. This is the case for Polish, which has few onsets and codas of five consonants, but a much larger variety of smaller clusters than what is shown here. In the next section, I will discuss issues of the prevalence of Highly Complex syllable patterns in more detail.

1.4.2 **Prevalence of Highly Complex syllable patterns within languages**

Here I attempt to characterize the prevalence of Highly Complex syllable patterns in the sample. First I examine restrictions on the combinations of consonants occurring in Highly Complex structures in each language. Then I present information on the relative frequency (either quantified or impressionistic) of these patterns as reported in the language descriptions. Together, these measures provide a rough diagnostic for the relative prevalence of the target syllable patterns within the Highly Complex languages of the sample.

The analysis of restrictions on consonant combinations presented below is based primarily on the patterns of the smaller Highly Complex structures in each language. This is because the point here is to characterize the prevalence of Highly Complex patterns in general, and not just the maximal patterns occurring in each language. The analysis of restrictions on consonant combinations relies on patterns explicitly reported by the author. In some cases, no explicit description of consonant combinations is given, and I rely on patterns gleaned from the available examples.

For each language, I have classified the Highly Complex patterns which occur into three categories based on their combinatorial restrictions: Severely Restricted, Relatively Restricted, and Relatively Free. Where a language has Highly Complex structures in both margins and the patterns are qualitatively different, I examine the onset and coda separately. In (27)-(29) I give the definition for each category and illustrative examples from the data. The raw number of potential consonant combinations in a language is, of course, a function of the number of consonants in its phoneme inventory. I have attempted to define these categories

1 Syllable structure patterns in sample

so that they do not refer to or depend heavily upon the size of the consonant inventory of the given language.

(27) **Severely Restricted:** Just a handful of (< 5) Highly Complex sequences occur, and/or every member of the sequence has specific restrictions.

a. **Wutung** (*Sko*; Papua New Guinea)

Restrictions on onsets of four consonants:

Only /hmbɫ/ occurs.⁷

e.g. *hmbɫie*

‘left hand’ (*Marmion 2010* : 69)

b. **Doyayo** (*Atlantic-Congo*; Cameroon)

Restrictions on codas of four consonants:

C₁: must be /b g m ŋ/ (/b g/ usually realized as [β ɣ] in clusters)

C₂: must be /l r n/

C₃: must be /d t/

C₄: must be /s z/

Additionally, C₃ and C₄ must match in voicing.

e.g. *deβrts*

‘be cut off for’

(*Wiering 1995*: 41–42)

(28) **Relatively Restricted:** There are general restrictions on the voicing, place, or manner of some or all members, and/or specific restrictions on one or two (but not all) members.

a. **Lezgian** (*Nakh-Daghestanian*; Russia, Azerbaijan)

Restrictions on onsets of three consonants:

C₁: voiceless obstruent

C₂: voiceless obstruent or /r/

C₃: voiceless obstruent or sonorant

e.g. *k^hstaχ*

‘spoiled child’ (*Haspelmath 1993* : 37)

b. **Passamaquoddy-Maliseet** (*Algonquian*; Canada, USA)

Restrictions on onsets of three consonants:

⁷The /h/ here appears to be a separate consonant segment and does not represent a modification of the phonation of the following nasal. *Marmion (2010* : 54) describes it as a segment which can optionally elide preceding sonorant consonants.

1.4 Properties of highly complex syllable structure

Apart from a few exceptions, triconsonantal onsets or codas are always of the form CsC.

e.g. *kspison*

‘belt’

(LeSourd1993: 121)

- (29) **Relatively Free:** There may be a few abstract restrictions on consonant combinations, and/or combinations are described by author as free or unrestricted.

- a. **Yakima Sahaptin** (*Sahaptian*; USA)

Restrictions on codas of three and four consonants:

Clusters of glottalized or labialized obstruents do not occur.

e.g. *χipχp tawq'χf*

‘cottonwood’ ‘kerchief, neck scarf’

(Hargus & Beavert 2002 : 270-1)

The distribution of the languages with respect to the three categories above is given in Table 1.14.

Table 1.14: Degree of restriction on consonant combinations in Highly Complex syllable patterns.

Severely Restricted	Relatively Restricted	Relatively Free
Alamblak	Albanian	Cocopa
Bench	Camsá	Itelmen
Doyayo	Georgian	Nuu-chah-nulth
Kunjen	Kabardian	Polish (<i>onsets</i>)
Menya	Lezgian	Tashlhiyt
Qawasqar (<i>codas</i>)	Mohawk	Thompson
Wutung	Passamaquoddy-	Yakima Sahaptin
	Maliseet	Yine
	Polish (<i>codas</i>)	
	Qawasqar (<i>onsets</i>)	
	Semai	
	Tehuelche	
	Tohono O'odham	

1 Syllable structure patterns in sample

There are two languages — Polish and Qawasqar — which have different degrees of restriction in their Highly Complex onset and coda patterns. Besides Qawasqar, there are six languages for which all Highly Complex patterns are severely restricted. Interestingly, in only one of these (Doyayo) are the severely restricted patterns associated with specific morphologically complex sequences; in the others, they occur within morphemes. Most often, languages have Highly Complex structures that are relatively restricted in their consonant combinations. Besides Polish and Qawasqar, there are ten languages which have this pattern. Finally, there are seven languages besides Polish which have relatively free consonant combinations in their Highly Complex structures. It is striking that the set of languages with relatively free patterns is similar in size to the set of languages with severely restricted patterns, given the general rarity of languages with Highly Complex syllable structure.

Below I present information on the frequency of Highly Complex structures in the languages of the sample. Frequency of syllable patterns is explicitly remarked upon for only 16 of the 25 languages in this category. Most often, reports are impressionistic in nature, but occasionally a researcher provides type frequency data for patterns in the syllable inventory, lexicon, or text. In Table 1.15 I note the nature of the frequency data given for each language. Note that not all of the patterns reported below are strictly Highly Complex patterns; authors often did not make a distinction between different kinds of triconsonantal clusters, for instance.

Comparing the relative frequency patterns in Table 1.15 to the combinatorial restriction patterns in Table 1.14, we find some correspondences between patterns which are not all that surprising. For example, it follows that Wutung, whose Highly Complex syllable patterns are restricted to a single four-consonant onset (27a), would also have a very low type frequency of this pattern in its syllable inventory. Similarly, it is expected that Georgian and Polish, both of which have larger clusters and fewer restrictions on consonant combinations, should have a higher type frequency of these patterns in their syllable inventories.⁸ The other kinds of frequency data — type frequency in the lexicon and in running text — also show this correspondence, with higher frequencies typically corresponding to languages with freer consonant combinations in their Highly Complex patterns. It should also be noted that frequency patterns are reported for all but one

⁸Mohawk presents an unexpected pattern, in that its cluster patterns are relatively restricted but it has a type frequency of Highly Complex clusters which is on par with that of Georgian and Polish. This is due to the very small consonant phoneme inventory of the language (ten consonants), which limits the overall size of the syllable inventory.

1.4 Properties of highly complex syllable structure

Table 1.15: Reported frequency of Highly Complex syllable patterns.
Emphasis my own in all quotations.

Language	Nature of frequency data	Reported frequency of HC patterns
Bench	<i>Impressionistic</i>	“Syllable patterns ending in [CCC] have thus a very limited actual occurrence ” (Rapold 2006 : 92)
Camsá	<i>Impressionistic</i>	“Consonant clusters are very common in Camsá. [...] Clusters of three consonants are not as common in the language as clusters of two.” (Howard 1967 : 81-4)
Cocopa	<i>Impressionistic</i>	“[i]t is quite common to find Cocopa words consisting of a single vowel preceded by several consonants.” (Bendixen 1980 : 1)
Georgian	<i>Type frequency in syllable inventory</i>	28/276 (10%) of onset patterns occurring stem-initially are HC (calculated from data provided in Butskhrikidze 2002 : 197-205).
	<i>Type frequency in text</i>	In an excerpt of descriptive prose, 24/550 (4.4%) of word-initial patterns and 7/559 (1.3%) of word-final patterns consist of three or more consonants (Vogt 1958 : 79-80).
Kabardian	<i>Impressionistic</i>	“Clusters consist of not more than three, and in the large majority of cases, of two consonants.” (Kuipers 1960 : 29)
Kunjen	<i>Type frequency in lexicon</i>	“VCCCC syllables occur only as the initial syllable of the word, and have been recorded in only twenty words .” (Sommer 1969 : 35)
Itelmen	<i>Impressionistic</i>	“The frequent occurrence of 41 complex consonant clusters is one of the most notable traits of Itelmen phonology.”

1 Syllable structure patterns in sample

Table 1.16: Reported frequency of Highly Complex syllable patterns.
Emphasis my own in all quotations.

Language	Nature of frequency data	Reported frequency of HC patterns
Mohawk	<i>Type frequency in syllable inventory</i>	6/43 (14%) of word-initial onsets and 3/25 (12%) of word-final codas are HC (calculated from data provided in Michelson 1988 : 12-13).
Polish	<i>Type frequency in syllable inventory</i>	64/426 (15%) of onset patterns occurring word-initially are HC, 18/141 (13%) of coda patterns occurring word-finally are HC (calculated from data provided in Bargielowna 1950).
Tashlhiyt	<i>Type frequency in text</i>	451/5700 (7.9%) of syntactic words in running text are composed of voiceless obstruents only (Ridouane 2008 : 328f).
Thompson	<i>Impressionistic</i>	“Sequences of six obstruents are not uncommon. ” (Thompson & Thompson 1992 : 25)
Tohono O’odham	<i>Impressionistic</i>	Morphological and phonological processes “yield a high frequency of complex moras and very intricate syllables” (Hill & Zepeda 1992 : 355)
Wutung	<i>Type frequency in syllable inventory</i>	1/40 (2.5%) of onset patterns are HC (calculated from data provided in Marmion 2010).
Yakima Sahaptin	<i>Type frequency in lexicon</i>	13/295 (4.4%) of underived nouns and adjectives have onsets of three or four Cs, 8/295 (3%) have codas of

language with relatively free consonant combinations (Nuu-chah-nulth). Though quantitative type frequency data isn't given for Cocopa, Itelmen, or Thompson, the authors make a point of mentioning the high frequency and commonplace nature of Highly Complex structures in these languages.

Combining the results of the analyses in this section and §1.4.1, we can identify two extreme patterns in the prevalence of Highly Complex patterns in the data. On one extreme, there is a group of languages for which Highly Complex structures are a minor pattern. These languages have Highly Complex structures at only one syllable/word margin. The structures consist of three or maximally four consonants which are severely restricted in their combination, and have relatively low type frequencies (Table 1.17). On the other extreme, there is a group of languages for which Highly Complex structures are a prevalent pattern. These languages have Highly Complex structures at both syllable/word margins. The structures may be more than four consonants in length, are relatively free in their combination, and have relatively high type frequencies (Table 1.18).

Table 1.17: Languages with minor Highly Complex patterns.

Language	Family	Region
Alamblak	<i>Sepik</i>	AUSTRALIA & NEW GUINEA
Bench	<i>Ta-Ne-Omoti</i>	AFRICA
Doyayo	<i>Atlantic-Congo</i>	AFRICA
Kunjen	<i>Pama-Nyungan</i>	AUSTRALIA & NEW GUINEA
Menya	<i>Angan</i>	AUSTRALIA & NEW GUINEA
Wutung	<i>Sko</i>	AUSTRALIA & NEW GUINEA

Over half of the languages in the Highly Complex portion of the sample have syllable patterns which are at one of these extremes. There are different areal distributions for the two groups of languages. The languages with Highly Complex syllable structure as a very minor pattern include all those from the Australia & New Guinea macro-area, as well as two languages from Africa. The languages with prevalent Highly Complex patterns are spoken in parts of Eurasia, North America, and the Atlas Mountain region of Africa; i.e., regions identified in Chapter ?? as being well-known for their complex syllable patterns. I will return to discussion of these patterns in §1.5.

1 Syllable structure patterns in sample

Table 1.18: Languages with prevalent Highly Complex patterns.

Language	Family	Region
Cocopa	<i>Cochimi-Yuman</i>	NORTH AMERICA
Georgian	<i>Kartvelian</i>	EURASIA
Itelmen	<i>Chukotko-Kamchatkan</i>	EURASIA
Polish	<i>Indo-European</i>	EURASIA
Tashlhiyt	<i>Afro-Asiatic</i>	AFRICA
Thompson	<i>Salishan</i>	NORTH AMERICA
Tohono O’odham	<i>Uto-Aztecan</i>	NORTH AMERICA
Yakima Sahaptin	<i>Sahaptian</i>	NORTH AMERICA

1.4.3 Acoustic and perceptual characteristics

Researchers often remark upon the phonetic characteristics of the long tautosyllabic clusters of obstruents which are characteristic of most languages with Highly Complex syllable structure. Descriptions typically note the presence of salient release or aspiration of stops, transitional vocalic elements between consonants at different places or with different manners of articulation, and lengthened consonant articulation for syllabic obstruents. These descriptions are relevant in the establishment of Highly Complex syllable structure as a language type which may have specific acoustic characteristics in addition to abstract phonological characteristics. It is also possible that clues to the development of Highly Complex syllable structure may be found in the acoustic and perceptual properties of these clusters. For example, it has been found that clusters resulting from historically recent processes of vowel syncope may retain traces of the previous vowel in the transitions between consonants (cf. [Chitoran & Babaliyeva 2007](#) for Lezgian).

Descriptions of the acoustic and perceptual characteristics are available for 18/25 of the languages in the Highly Complex portion of the sample. This is somewhat remarkable, given that many of the languages are underdescribed, and that such detailed phonetic descriptions of consonant clusters are not a standard topic for inclusion in language references. These descriptions are presented in [Table 1.19](#).

Even though many of the descriptions make mention of ‘epenthetic’ vowels, the patterns described above are consistent with those features listed by [Hall \(2006\)](#) as being associated with intrusive vowels. The transitional elements in

1.4 Properties of highly complex syllable structure

Table 1.19: Descriptions of acoustic and perceptual characteristics of clusters in languages with Highly Complex syllable structure. Languages omitted due to lack of description are Bench, Doyayo, Kabardian, Passamaquoddy-Maliseet, Polish, Qawasqar, and Wutung.

Language	Description of phonetic realization of consonant clusters
Alamblak	“Open transition,” transcribed as [i], varies freely with release in obstruent and other sequences (Bruce 1984 : 56-9).
Albanian	Release between obstruents varies freely with much rarer epenthetic [ə] in slow or careful speech (Klippenstein 2010 : 24-6).
Camsá	“Nonphonemic transitional vocoid [ə]” occurs between stops or consonant plus nasal at different points of articulation; initial fricatives are lengthened and may have voiceless or voiced off-glide, transcribed as [ʰ] or [ʰ̥], before a non-fricative consonant at a different place of articulation (Howard 1967 : 81).
*Cocopa	Consonants in some sequences separated by “anaptyctic phonetic vowel” or “indistinct ‘murmur’ vowel” whose quality, transcribed [i], [a], or [u], is determined by surrounding consonants (Crawford Jr 1966 : 37-45).
Georgian	Stops in sequences nearly always released, sometimes with voicing if both are voiced; voiceless stops and affricates have strongly aspirated release; length of interval between C ₁ and C ₂ release depends on relative place of articulation of the consonants (Chitoran 1999).
Itelmen	Indeterminant “overtone” transcribed as [ə] and described as “extremely short, with an overtone indeterminant in timbre,” occurs in words without vowels and certain consonant combinations (Volodin 1976 : 40-1; translation SME).
Kunjen	“Brief transitional vocoids” may sometimes be heard between consonants in a cluster. (Sommer 1969 : 33)
Lezgian	Before a voiceless stop or fricative, voiceless stops are always aspirated (Haspelmath 1993 : 47); in clusters resulting from historical or synchronic syncope, traces of previous vowel remain audible in stop release and fricative noise (Chitoran & Babaliyeva 2007).
Menya	“Non-homorganic consonants are phonetically separated by extremely short vocalic segments which are more and more not being written”; quality of short segments is conditioned by surrounding consonants and vowels (Whitehead 2004 : 9, 226).
Mohawk	Stops are “strongly aspirated” before another

Table 1.20: Descriptions of acoustic and perceptual characteristics of clusters in languages with Highly Complex syllable structure. Languages omitted due to lack of description are Bench, Doyayo, Kabardian, Passamaquoddy-Maliseet, Polish, Qawasqar, and Wutung. * indicates that reported pattern is for syllables with obstruent nuclei.

Language	Description of phonetic realization of consonant clusters
*Semai	Minor syllables consisting of consonants are “clearly heard and perceived as distinct syllables.” (Sloan 1988 : 321). Vocalic element in consonantal minor syllable “usually a very short, non-phonemic, epenthetic [ə]”, but can vary in quality, and is “optional if the two consonants are easily pronounced without the epenthetic vowel.” (Philips2007)
*Tashlhiyt	Short “voiced transitional vocoids” whose quality is predictable by surrounding vowels split consonant sequences when one is voiced (Dell & Elmedlaoui 2002 : 16); Gordon & Nafi (2012 : 16) report this for occasional sequences of voiceless consonants. Ridouane (2008 : 210) reports that “stop release is obligatory before another stop which is not homorganic with it” and Grice et al. (2015) find that the “vocoid” is not entirely predictable from the voicing properties of surrounding Cs and that its presence is partly conditioned by intonational prominence.
*Tehuelche	The “accumulation of consonants is made possible by the development of [...] supporting vowels.” These have “a neutral vowel quality which play the role of lubricator and which corresponds to the neutralization of all other vowels,” and are transcribed as [ə] or [ʊ] depending upon consonantal environment (FernándezGarayHernández2006: 13; FernándezGaray1998; translation RNS)
Thompson	Plain stops are “somewhat aspirated” before another stop and often before spirants, and strongly aspirated syllable-finally (Thompson & Thompson 1992 : 4). “Laryngeals are usually separated from preceding obstruents by a brief central vowel” whose precise quality is determined by the consonantal environment (1992: 44).
Tohono O’odham	Surface clusters resulting from historical vowel deletion have “very short, voiceless elements”, phonetically transcribed as [h] but which may retain previous vowel quality coloration in the case of high vowels (Hill & Zepeda 1992 : 356). Combinations of voiceless stops

these clusters are characterized by neutral vowel qualities that may be heavily influenced by surrounding consonants, and may vary in their duration and voicing. In some cases the transitions are described as occurring between specific combinations of consonants with different places of articulation.

Most of the references consulted for the above analysis were written by researchers who are not native speakers, and for whom the different timing patterns in consonant sequences in these languages may be especially salient.⁹ As mentioned in §1.2.2, native speakers are often unaware of the presence of these transitional elements, and when they are aware of them, view them as optional. Menya provides an interesting illustration of this in its writing conventions. The short vocalic elements between non-homorganic consonants in the language are written only sporadically by literate native speakers, and when they are written, there is unsystematic variation in the grapheme used (Whitehead 2004 : 9, 226; the quote given in Table 1.19 may also be suggestive of a recent process of vowel reduction). Another piece of evidence for determining the intrusive nature of a vowel is in its ‘invisibility’ to phonological processes. In some cases, explicit descriptions of this are given. For example, the vocalic element transcribed as [ɪ] that occurs between nasals and back stops or affricates in Nuuchahnulth is explicitly described as not being included in the syllable count which determines a vowel lengthening pattern in the language (Rose 1981 : 27).

The striking similarities in the phonetic descriptions of Highly Complex structures in the 18 languages in Table ?? indicate that the languages of this category share more in common than just phonological structure. The presence of transitional elements is a prominent phonetic characteristic of Highly Complex syllable structure. It is also notable that the phonetic descriptions of syllabic obstruents in Cocopa, Semai, Tashlhiyt, and Tehuelche are virtually indistinguishable from the phonetic descriptions of tautosyllabic clusters in the other languages. This recalls Hall’s observation that vowel intrusion and syllabic consonants are motivated by similar processes of gestural overlap, and is further justification for grouping these languages together with the others.

1.5 Discussion

As mentioned in §1.1, the studies in this chapter serve two purposes: first, to provide a baseline characterization of syllable patterns in the language sample as

⁹It is interesting to note that for Polish, which has a wealth of available descriptive material written in English by native speakers of Polish, I could find few details on the phonetic characteristics of clusters.

1 Syllable structure patterns in sample

a whole; and second, to elucidate in greater detail the specific patterns occurring in languages with Highly Complex syllable structure.

In Table 1.21 I summarize the results from §1.3 regarding syllable patterns in the language sample as a whole, and describe how the findings relate to syllable structure complexity. An asterisk (*) indicates that the pattern was found to be statistically significant.

Table 1.21: Summary of findings regarding syllable patterns in language sample as a whole.

Aspect of syllable structure	Finding
1. <i>Relationship between maximal onset and coda complexity (§1.3.2)</i>	A large cluster at one margin typically implies a large cluster at the other margin.
2.* <i>Obligatoriness of syllable margins (§1.3.3)</i>	Least common in S lgs, most common in HC languages.
3. <i>Complex vocalic nuclei (§1.3.4)</i>	Much less likely to occur in S languages.
4.* <i>Presence of syllabic consonants (§1.3.5)</i>	Least common in S lgs, most common in HC languages.
5. <i>Morphological constituency patterns (§1.3.6)</i>	
a.* <i>of maximal syllable margins</i>	Morphologically complex patterns increase with syllable structure complexity.
b. <i>of syllabic consonants</i>	More likely to be found in grammatical items as syllable structure increases.

Some of the analyses presented in §1.3, corresponding to the findings in lines 1-3 of Table 1.21, were exploratory and conducted without any underlying hypotheses. Nevertheless these analyses yielded relevant results with respect to syllable structure complexity.

The relationship between maximal onset and coda complexity, in which the presence of large (four or more consonants) sequences at one syllable margin in a language typically implies the presence of large sequences at the other margin, is especially interesting. This is not an expected pattern in terms of probabilistic distribution: if onsets and codas are independent structures, then we would ex-

pect to observe a wider range of combinations in maximal onset and coda sizes. However, if syllable structure is viewed not as an entity with abstract phonological motivations, but as a phenomenon reflecting articulatory routines carried out over many generations in the history of a language, this pattern may not be surprising. It is reasonable to imagine, for instance, a scenario in which a strong tendency toward vowel reduction in a language with prefixation and suffixation on stress-carrying stems might result in the eventual complete gestural overlap and deletion of many or most unstressed vowels, yielding long clusters of consonants in both word-marginal contexts. This issue will be discussed further in Chapters 5, 6, and 8.

Similarly, the pattern observed with respect to obligatory syllable margins and syllable structure complexity is not necessarily expected. In languages with Simple syllable structure, obligatory onsets significantly limit the size of the syllable inventory. However, this effect on the syllable inventory would be much smaller in languages in the other syllable structure complexity categories. It does not necessarily follow that the Highly Complex category should have a higher rate of obligatory margins than the other categories. Three of the languages with obligatory onsets in the Highly Complex category (Thompson, Tohono O'odham, and Yakima Sahaptin), as well as one language in this group which likely had obligatory onsets historically (Itelmen), also happen to be in the group of languages whose Highly Complex patterns are most prevalent (§1.4.2). Placed in the context of languages in which large consonant clusters are prevalent and the consonant to vowel ratio in speech is presumably higher than average, perhaps the high rate of obligatory syllable margins is not surprising for this category.

Greater syllable structure complexity is also associated with a greater likelihood of a language having complex vocalic nuclei (long vowels, diphthongs, or vowel sequences) and/or syllabic consonants. The specific combinatorial restrictions between syllable margins and nuclei are not explored or quantified here. However, the general effect of these patterns could be that as syllable structure complexity increases, so does the range of possible syllable types, and not just as a result of increased consonant combinations. The diversity in syllable margins which provide the basis for the definitions of syllable structure complexity may be accompanied by greater diversity in syllable nuclei as syllable structure complexity increases. This issue will be explored in greater depth in §??.

The data here confirms the hypothesis that languages with more complex syllable structure are more likely to have syllabic consonants. This pattern is strongly present even when syllabic obstruents are not considered. These results are not in line with the predictions of *Isačenko (1939)*, who observed that 'vocalic' lan-

1 Syllable structure patterns in sample

languages — a term defined in part by a low consonant to vowel ratio and simpler syllable patterns — are more likely to develop syllabic sonorants. Isačenko's study is limited to Slavic languages, so it is possible that this group of languages presents an exception to the crosslinguistic patterns. However, the findings of the current study should be considered in the context of Isačenko's larger point that 'vocalic' and 'consonantal' languages are characterized not only by different segmental inventory and syllable patterns, but also by different diachronic and synchronic processes of language change. It is known that syllabic consonants often come about through vowel reduction (Bell 1978), and indeed we observe a set of variable syllabic consonants in the data which come about through vowel reduction. Independently of the observation that vowel reduction is known to create tautosyllabic clusters in some languages, the syllabic consonant patterns here suggest a higher occurrence of vowel reduction processes, both diachronic and synchronic, in languages with more complex syllable structure. This issue will be explored in greater detail in Chapter ??.

The hypothesis with respect to morphological patterns in the data was also confirmed. As syllable structure increases, maximal onset and coda clusters are more likely to exhibit morphologically complex patterns, and syllabic consonant inventories are more likely to have members which correspond to grammatical morphemes. These findings point toward a strong influence of morphology in the emergence of more complex, and specifically Highly Complex, syllable structures. The studies here are largely limited to phonological systems and do not allow for an in-depth study of the morphological patterns of the language sample. However, the issue of the role of morphology in the syllable patterns of languages in the Highly Complex group will be revisited in Chapter ??.

The analysis of the Highly Complex patterns in §1.4 reveals important patterns within this group of languages that should be considered in the coming chapters. The first is that there are measures by which this is not a coherent group of languages. Analyses of the specific syllable structures occurring in these languages, as well as the restrictions on these structures and their frequency of occurrence, suggests that there are instead several groups. In one group of six languages from Africa and Australia & New Guinea, Highly Complex syllable structure is an extremely minor pattern, and includes low frequencies of highly restricted structures, often containing several sonorants, at one syllable margin. In another group of eight languages, mostly from Eurasia and North America, Highly Complex syllable structure is a prevalent pattern, and involves high frequencies of long, fairly unrestricted strings of obstruents at both syllable margins. While it wasn't explicitly discussed in §1.4, these two groups also have different mor-

phological patterns in their maximal syllable structures. Most of the languages with minor Highly Complex patterns have morpheme-internal patterns for their maximal syllable margins (the two exceptions being Bench and Doyayo). By comparison, all of the languages with prevalent Highly Complex patterns have morphologically complex patterns for their maximal syllable/word margins.

Thus there are two extreme groups within the Highly Complex category which can be set apart from the rest on the basis of having different sets of coherent behavior in their syllable patterns. The other 11 languages of the sample fall somewhere between these two extremes. In the upcoming studies it will be discussed how languages on the two extremes of the Highly Complex category also exhibit coherent differences in their segmental inventories, stress patterns, and phonological processes, and how the languages in between the extremes behave more like one group or another.

The second important finding in §1.4 is that there is another measure by which the languages of the Highly Complex category *are* a coherent group of languages. In all languages for which phonetic properties of clusters were described, Highly Complex clusters were described as having largely similar acoustic and perceptual characteristics. This is true of both languages with large tautosyllabic clusters and those with syllabic obstruents, suggesting that these phenomena are qualitatively similar and/or have similar origins. This point will be revisited in Chapter ??.

References

- Ameka, Felix K. 1991. *Ewe: Its grammatical constructions and illucutionary devices*. Australian National University dissertation.
- Anderson, Victoria. 2000. Giving weight to phonetic principles: University of California, Los Angeles dissertation.
- Bagemihl, Bruce. 1991. Syllable structure in Bella Coola. *Linguistic Inquiry* 22(4). 589–646.
- Bargiełowna, M. 1950. Grupy fonemów spółgłoskowych współczesnej polszczyzny kulturalnej. *Biuletyn Polskiego Towarzystwa Językoznawczego* 10. 1–25.
- Bell, Alan. 1978. Syllabic consonants. In Charles. A. Ferguson Greenberg Joseph H. & Edith A. Moravcsik (eds.), *Universals of human language*, vol. 2, 153–201. Stanford: Stanford University Press.
- Bendixen, Birgitte Holt. 1980. *Phonological and temporal properties of cocopa*. University of California San Diego dissertation.
- Blevins, Juliette. 1995. The syllable in phonological theory. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 206–44. Oxford: Basil Blackwell.
- Blevins, Juliette. 2006. Syllable typology. In Keith Brown (ed.), *Elsevier encyclopedia of language and linguistics. Second edition*, vol. 12, 333–337. Oxford: Elsevier.
- Bonvillain, Nancy. 1973. *A grammar of akwesasne mohawk*. Ottawa: National Museum of Man.
- Breen, Gavan & Rob Pensalfini. 1999. Arrernte: A language with no syllable onsets. *Linguistic Inquiry* 30(1). 1–25.
- Browman, Catherine & Louis Goldstein. 1992. “targetless” schwa: An articulatory analysis. In Gerard Docherty & D. Robert Ladd (eds.), *Papers in laboratory phonology II: Gesture, segment, prosody*, 26–56. Cambridge: Cambridge University Press.
- Bruce, Leslie P. 1984. *The alamblak language of Papua New Guinea (East sepik) (Pacific linguistics, series c, No. 81)*. Canberra: Research School of Pacific Studies, Australian National University.

References

- Butskhrikidze, Marika. 2002. *The consonant phonotactics of Georgian*. Universiteit Leiden dissertation.
- Campbell, Lyle. 1974. Theoretical implications of kekchi phonology. *International Journal of American Linguistics* 40. 269–278.
- Chang, Anna Hsiou-chuan. 2006. *A reference grammar of Paiwan*. The Australian National University dissertation.
- Chitoran, Ioana. 1999. Accounting for sonority violations: The case of Georgian consonant sequencing. *Proceedings of the XIV International Congress of Phonetic Sciences, San Francisco 1999*. 101–104.
- Chitoran, Ioana & Ayten Babaliyeva. 2007. *An acoustic description of high vowel syncope in Lezgian*. International Congress of Phonetic Sciences XVI: 2153–56.
- Côté, Marie-Helene. 2011. Final consonants. In Marc Van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, no pages. no address: Wiley-Blackwell.
- Crawford Jr, James Mack. 1966. *The cocopa language*. University of California at Berkeley dissertation.
- Davidson, Matthew. 2002. *Studies in Southern wakashan (nootkan) grammar*. State University of New York at Buffalo dissertation.
- Davis, Stuart & Karen Baertsch. 2011. On the relationship between codas and onset clusters. In Charles Cairns & Eric Raimy (eds.), *Handbook of the syllable*, 71–97. Leiden: Brill.
- Dell, François & Mohamed Elmedlaoui. 2002. *Syllables in tashlhiyt berber and in moroccan Arabic*. Dordrecht: Kluwer.
- Diffloth, Gérard. 1976. *Expressives in semai*. Oceanic Linguistics Special Publications 13, Austroasiatic Studies Part I: 249–264.
- Dixon, R. M. W. 1970. Olgolo syllable structure and what they are doing about it. *Linguistic Inquiry* 1(2). 273–276.
- Dol, Philomena. 2007. *A grammar of maybrat: A language of the bird's head peninsula, Papua province, Indonesia*. Canberra: Pacific Linguistics.
- Donohue, Mark, Rebecca Hetherington, James McElvenny & Virginia Dawson. 2013. *World phonotactics database*. <http://phonotactics.anu.edu.au>, accessed 2016-10-16. Department of Linguistics, The Australian National University.
- Dressler, Wolfgang U. & Katarzyna Dziubalska-Kołaczyk. 2006. Proposing morphonotactics. *Rivista di Linguistica* 18(2). 249–266.
- Dressler, Wolfgang U., Katarzyna Dziubalska-Kołaczyk & Lina Pestal. 2010. Change and variation in morphonotactics. *Folia Linguistica* 31. 51–67.

- Dressler, Wolfgang U., Miroslava Hliničanová, Matej Ďurčo, Karlheinz Mörtz & Katharina Korecky-Kröll. 2015. Phonotactic vs. *morphonotactic obstruent clusters in Slovak and German*. *Rivista di Linguistica* 27(1). 45–60.
- Eberhard, David Mark. 2009. *Mamaindê grammar*: Vrije Universiteit, Amsterdam dissertation.
- Fenk, August & Gertraud Fenk-Oczlon. 1993. Menzerath's law and the constant flow of linguistic information. In Reinhard Köhler & Burghard B. Rieger (eds.), *Contributions to quantitative linguistics: Proceedings of the first international conference on quantitative linguistics, QUALICO, trier, 1991*. Dordrecht: Kluwer.
- Filchenko, Andrey Yury. 2007. *A grammar of Eastern khanty*. Rice University dissertation.
- Gasser, Emily & Claire Bowern. 2014. Revisiting phonological generalizations in Australian languages. In Kingston John, Claire Moore-Cantwell, Joe Pater & Robert Staubs (eds.), *Supplemental proceedings of phonology 2013*. Washington DC: Linguistic Society of America.
- Georg, Stefan & Alexander P. Volodin. 1999. *Die itelmenische Sprache: Grammatik und texte*. Weisbaden: Harrassowitz Verlag.
- Gerner, Mattias. 2013. *A grammar of nuosu*. Berlin/Boston: De Gruyter Mouton.
- Givón, Talmy. 2011. *Ute reference grammar*. Amsterdam/Philadelphia: John Benjamins Publishing.
- Goldstein, Louis, Ioana Chitoran & Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In Proceedings of the XVIth International Congress of Phonetic Sciences, 6-10 August 2007, Saarbrücken, Germany. 241-244.
- Gordon, Matthew K. 2006. *Syllable weight: Phonetics, phonology, typology*. New York: Routledge.
- Gordon, Matthew K. 2016. *Phonological typology*. Oxford: Oxford University Press.
- Gordon, Matthew & Latifa Nafi. 2012. Acoustic correlates of stress and pitch accent in tashlhiyt berber. *Journal of Phonetics* 40(5). 706–724.
- Greenberg, Joseph. 1965. Some generalizations concerning initial and final consonant sequences. *Linguistics* 18. 5–34. Reprinted in 1978 in Greenberg, Joseph H. (ed.), *Universals of human language*, vol. 2: Phonology. Stanford, CA: Stanford University Press.
- Grice, Martine, Rachid Ridouane & Timo B. Roettger. 2015. Tonal association in tashlhiyt berber: Evidence from polar questions and contrastive statements. *Phonology* 32. 241–266.

References

- Guillaume, Antoine. 2008. *A grammar of cavineña*. (Mouton Grammar Library, 44.) Berlin: Mouton de Gruyter.
- Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23(3). 387–429.
- Hanson, Rebecca. 2010. *A grammar of Yine (piro)*. La Trobe University, Melbourne dissertation.
- Hargus, Sharon & Virginia Beavert. 2002. Yakima sahapitin clusters and epenthetic [i]. *Anthropological Linguistics* 44(3). 231–277.
- Hargus, Sharon & Virginia Beavert. 2006. Word-initial clusters and minimality in Yakima sahapitin. *Phonology* 23. 21–58.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. (Mouton Grammar Library, 9.) Berlin: Mouton de Gruyter.
- Hill, Jane H. & Ofelia Zepeda. 1992. Derived words in tohono o’odham. *International Journal of American Linguistics* 58(4). 355–404.
- Hoard, James E. 1978. Syllabification in Northwest Indian languages, with remarks on the nature of syllabic stops and affricates. In Alan Bell & Joan Bybee Hooper (eds.), *Syllables and segments*, 59–72. New York: North-Holland Publishing Company.
- Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
- Howard, Linda. 1967. Camsa phonology. In Viola G. Waterhouse (ed.), *Phonemic systems of Colombian languages*, 73–87. Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Isačenko, A. V. 1939. Versuch einer typologie der slavischen sprachen. *Linguistica Slovaca* 1. 64–76.
- Jassem, Wiktor. 2003. Polish. *Journal of the International Phonetic Association* 33(1). 103–107.
- Kay, Gillian. 1995. English loanwords in Japanese. *World Englishes* 14(1). 67–76.
- Kim, Eun-Sook. 2003. *Theoretical issues in Nuu-chah-nulth phonology and morphology*. University of British Columbia dissertation.
- Klippenstein, Rachel. 2010. Word-initial consonant clusters in albanian. *Ohio State University Working Papers in Linguistics* 59. 10–32.
- Kreitman, Rina. 2008. *The phonetics & phonology of onset clusters*: Cornell University dissertation.
- Kuipers, Aert H. 1960. *Phoneme and morpheme in kabardian*. The Hague: Mouton.
- Kwon, Harim, Ioana Chitoran, Marianne Pouplier, Tomas Lentz & Philip Hoole. 2017. *Cross-linguistic differences in the perception of articulatory timing lag in*

- onset clusters*. Paper presented at 91st Annual Meeting of the Linguistic Society of America. 5-8 January 2017, Austin, USA.
- Lacrampe, Sébastien. 2014. Lelepa: Australian National University dissertation.
- Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand & Bernard Lak (eds.), *Current trends in phonology. Models and methods*, 419–443. Paris: CNRS, ESRI.
- Maddieson, Ian. 2013. Syllable structure. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Maddieson, Ian, Sébastien Flavier, Egidio Marsico, Christophe Coupé & François Pellegrino. 2013. LAPSyD: Lyon-Albuquerque Phonological Systems Database. In Proceedings of the 14th Interspeech Conference, Lyon, France, 25-29 August 2013.
- Maddieson, Ian & Kristin Precoda. 1992. Syllable structure and phonetic models. *Phonology* 9. 45–60.
- Marmion, Douglas. 2010. *Topics in the phonology and morphology of Wutung*. The Australian National University dissertation.
- Mason, J. Alden. 1950. *The language of the papago of Arizona*. Philadelphia: American Philosophical Society.
- Matteson, Esther. 1965. *The piro (arawakan) language*. Berkeley & Los Angeles: University of California Press.
- Matteson, Esther & Kenneth L. Pike. 1958. *Non-phonemic transition vocoids in Piro (Arawak) Miscellanea Phonetica* III. 22–30. . London: International Phonetic Association.
- Michelson, Karin. 1988. *A comparative study of Lake-Iroquoian accent*. Dordrecht: Kluwer.
- Minthorn, Modesta. 2005. *Sonority and the Umatilla syllable*. (M.A. thesis, University of Oregon).
- Morelli, Frida. 1999. *The phonotactics and phonology of obstruent clusters of optimality theory*. University of Maryland dissertation.
- Morelli, Frida. 2003. The relative harmony of /s+stop/ onsets: Obstruent clusters and the sonority sequencing principle. In Caroline Féry & Ruben van de Vijver (eds.), *The syllable in optimality theory*, 356–372. Cambridge: Cambridge University Press.
- Murray, Robert W. & Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59(3). 514–528.
- Napoleão de Souza, Ricardo. 2019. The interaction of domain-initial effects with lexical stress: University of New Mexico dissertation.

References

- Olson, Kenneth. 2003. *Word minimality and non-surface-apparent opacity in Mono*. Talk given at WOCAL 4, Rutgers University.
- Orzechowska, Paula. 2012. Effects of morphology on the syllable structure. In Thomas Stolz, Nicole Nau & Cornelia Stroh (eds.), *Monosyllables: From phonology to typology*, 91–114. Berlin: Walter de Gruyter.
- Pan, Chia-jung. 2012. *A grammar of lha'alua, an Austronesian language of Taiwan*. (Doctoral dissertation, James Cook University.)
- Polian, Gilles. 2006. *Éléments de grammaire du tseltal, une langue maya du mexique*. Paris: L'Harmattan.
- Rapold, Christian. 2006. *Towards a grammar of benchnon*. Rijksuniversiteit te Leiden dissertation.
- Ridouane, Rachid. 2008. Syllables without vowels: Phonetic and phonological evidence from tashlhiyt berber. *Phonology* 25. 321–359.
- Rose, Suzanne M. 1981. *Kyuquot grammar*. University of Victoria dissertation.
- Rousset, Isabelle. 2004. *Structures syllabiques et lexicales des langues du monde*. Université Grenoble dissertation.
- Sloan, Kelly. 1988. Bare-consonant reduplication: Implication for a prosodic theory of reduplication. In Hagit Borer (ed.), *Proceedings of the seventh West coast conference on formal linguistics*, 310–330. Stanford: Stanford Linguistics Association.
- Sommer, Bruce A. 1969. *Kunjen phonology: Synchronic and diachronic*. (Pacific Linguistics, Series B, 11.) Canberra: Australian National University.
- Sommer, Bruce A. 1970. *Kunjen Syntax: A Generative View*. (Doctoral dissertation, University of Hawai'i at Mānoa.)
- Sommer, Bruce A. 1981. The shape of kunjen syllables. In D. L. Goyvaerts (ed.), *Phonology in the 1980's*, 231–244. Gent: E. Story-Scientia.
- Stirtz, Timothy M. 2011. *A grammar of gaahmg: A nilo-Saharan language of Sudan*. Utrecht: LOT.
- Thompson, Laurence C. & M. Terry Thompson. 1992. *The Thompson language*. (University of Montana Occasional Papers in Linguistics, 8.) Missoula, Montana: Linguistics Laboratory, University of Montana.
- Vallée, Nathalie, Solange Rossato & Isabelle Rousset. 2009. Favoured syllabic patterns in the world's languages and sensorimotor constraints. In François Pellegrino, Egidio Marsico, Ioana Chitoran & Christophe Coupé (eds.), *Approaches to phonological complexity*, 111–139. New York: Walter de Gruyter.
- VanDam, Mark. 2004. Word final coda typology. *Journal of Universal Language* 5. 119–148.

- Vaux, Bert & Andrew Wolfe. 2009. The appendix. In Eric Raimy & Charles Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 101–143. Cambridge: MIT Press.
- Vogt, Hans. 1958. Structure phonémique du géorgien: Etude sur le classement des phonèmes et des groupes de phonèmes. *Norsk Tidsskrift for Sprogvidenskap* 18. 5–90.
- Volodin, Aleksandr P. 1976. *Itel'menskij jazyk*. Leningrad: Nauka.
- Whitehead, Carl R. 2004. *A reference grammar of Menya, an angan language of Papua New Guinea*. Ann Arbor: UMI.
- Yumitani, Yukihiro. 1998. *A phonology and morphology of jemez towa*. (Ph.D. dissertation: University of Kansas).
- Zec, Draga. 2007. The syllable. In Paul De Lacy (ed.), *The Cambridge handbook of phonology*, 161–194. Cambridge: Cambridge University Press.

Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit <http://www.langsci-press.org/donate> to provide financial support or register as a community proofreader or typesetter at <http://www.langsci-press.org/register>.



Highly complex syllable structure

Set blurb on back with \BackBody{my blurb}

