Spring 5-21-2021

A Phonological Analysis of the Word-Borrowing Process in Volapük

Yutong Zhang

WASHINGTON UNIVERSITY IN ST. LOUIS

Linguistics Program

A Phonological Analysis of the Word-Borrowing Process in Volapük

by

Yutong Zhang

A thesis presented to

the Linguistics Program

of Washington University in

partial fulfillment of the

requirements for the degree

of Bachelor of Arts

May 2021

St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Abbreviations

OT Optimality Theory

RCD Recursive Constraint Demotion

SSP Sonority Sequencing Principle

ABSTRACT OF THE THESIS

A Phonological Analysis of the Word-Borrowing Process in Volapük

by

Yutong Zhang

Bachelor of Arts in Linguistics

Washington University in St. Louis, 2021

This paper analyzes the phonological process present in the word-borrowing process in *Volapük nulik*, the later version of Volapük – one of the world's first constructed international auxiliary languages that achieved more than a million speakers – through an analogy with the loanword adaptation process taking place in a natural language. It examines the emergent phonological patterns within this process, despite the inherited arbitrariness of any constructed languages, and compares them with the prescriptive rules regulating the word-borrowing process featured in its grammar. The paper is divided into three parts: Part I generalizes the syllable shape in Volapük from a 1000-base sample wordlist and compares it in general with the prescriptivism of De Jong (1931); Part II provides a mandatory syllable template from the result in Part I; Part III then explains the analogous loanword adaptation process in Volapük further with an approach of Optimality Theory and offers ranking information of a variety of constraints deemed relevant in the process. The result of the descriptive analysis not only indicates a consistency between the descriptive data and the prescriptive rules, i.e., all prescriptive rules are proven to be applicable and mandatory, but also reveals

additional syllable patterns that are not explicitly mentioned in its prescriptive grammar. Furthermore, OT analysis indicates that, despite a contradiction due to the arbitrary nature of the language, constraints relevant to the formation of the mandatory syllable shape are generally ranked higher than other faithfulness constraints, which suggests a preference for maintaining the prescribed syllable structure in the analogous loanword adaptation process in Volapük.

1. INTRODUCTION. Volapük is the first a posteriori constructed international auxiliary language achieving more than a million speakers prior to the rise of Esperanto. Constructed in late 1870s by the German priest Johann Martin Schleyer, it underwent constant modifications, both grammatical and lexical: for instance, a stem can have various forms in different Volapük versions. The earlier versions constructed by Schleyer before 1900s are, as a whole, known as *Volapük rigik* 'original Volapük', while the current version, *Volapük nulik* 'new Volapük', was a revised version by the Dutch doctor Arie de Jong in the 1930s on the basis of Schleyer's late versions. At present, there are less than a dozen Volapük speakers throughout the world who speak the language fluently, most of whom belong to *Kadäm Volapüka* 'The Volapük Academy', which has been administrating the language movement and regulating the use of the language since Schleyer's era. The Academy also publishes a monthly gazette, *Vög Volapüka* 'The Voice of Volapük', in which a variety of Volapük texts are published.

In terms of morphology, Volapük behaves like an agglutinative language in which a wide variety of inflectional and derivational affixes is present. Most of the monomorphemic bases also serve as free morphemes: the majority of them are nominal (e.g., ‹drim› 'dream'), to which inflectional (e.g., ‹drim-s› 'dreams') and derivational affixes (e.g., ‹drim-an› 'dreamer') can be attached, whereas others are function words (e.g., ‹e› 'and') and monomorphemic adverbs (e.g., ‹ba› 'probably'). In some rarer cases, however, the monomorphemic bases are bound and must first be suffixed by a corresponding inseparable derivational suffix to form a free morpheme that has meaning. Most frequently, these suffixes denote time (e.g., ‹mäz-ul› 'March'), directions (e.g., ‹sul-üd› 'the south'), interjections (e.g., ‹n-ö› 'no'), and several adverbs (e.g., ‹ib-o› 'namely, indeed') and prepositions (e.g., ‹v-ü› 'between'). Additionally, verbs and adjectives are typically derived by attaching a derivational suffix to a free nominal base and undergo further inflectional process via affixation. It is also possible to build compound words by binding two or more free morphemes together, often linked by ‹-a›, the genitive suffix: ‹vol-a-pük› 'universal language, language of the world'. The robust affixation in Volapük helps to better understand the prescriptive rules as described in Section 4.1.

My research aims at analyzing the word-borrowing process in Volapük nulik by studying 1000 word-bases. Understandably, the process in constructed languages such as Volapük behaves inherently differently from that in any native languages, as the former is usually regarded as more conscious and perhaps arbitrary, namely, inclined more to prescriptive rules rather than possessing emergent and systematic phonotactics, due to the artificial nature of these languages. By adopting similar theoretical tools used for natural languages, however, this thesis attempts to look for potential emergent patterns in the word-borrowing process of a constructed language, namely Volapük, which might not be explicitly mentioned or explained in its prescriptive grammar, meanwhile also acknowledging the certain presence of arbitrariness in this process (e.g., the segmental constraints discussed in Section 6.2.3).

The research is divided chiefly into three parts: Part I (Section 4) analyzes the overall syllable shape in Volapük from the data and compares it with the prescriptive rules; Part II (Section 5) provides a mandatory syllable template generalized from the result in Part I; Part III (Section 6) then examines the data further with an optimality-theoretic approach analogous to the loanword adaptation model in natural languages and offers ranking information of a variety of constraints deemed relevant in the loanword curation process of Volapük. In spite of this analogy, crucial differences do arise in the case of Volapük compared to pure loanword adaptation as observed in a natural language, which is further discussed in Section 6.1. Overall, the process seems so rigid that the vast majority of Volapük bases (approximately 97%) fit into the mandatory syllable template, and this raises the question if there is a consistent process by which the source words are borrowed into the prosodically restricted Volapük bases in terms of insertions, deletions, and segmental changes.

2. METHODS.

2.1. DATA COLLECTION. For the purpose of this research, a list of all Volapük nulik bases (or, in the case of bound bases, the smallest meaningful elements) present in the text of the Volapük gazette, *Volapükagased pro Nedänapükans* 'Volapük Gazette for Dutch Speakers', from the year 1932 to 1935 has been extracted as the sample data to a spreadsheet along with their prescriptive pronunciation in IPA, their corresponding translation in English, their possible source language (marked according to the ISO 639-3 standard for language abbreviations), and their counterpart in source language (i.e., the etymology of the bases) as well as the pronunciation in

IPA of these corresponding source items. The final list yields a total of exactly 1,000 bases present in the text of the gazette during this time period (out of all existing 6,711 Volapük nulik bases), while those bases known to have been invented entirely from scratch – primarily grammatical words such as pronouns and numerals – are excluded from the list. The spreadsheet "Supplement1" is attached to the thesis to provide more detail.

The plausible etymology of a Volapük base on the wordlist is determined with the assistance from my previously compiled Volapük etymological dictionary (Zhang 2018) and, in a more general way, according to the description of the word-borrowing process in Volapük by De Jong (1931), namely, reflected by a similarity in either pronunciation or orthography between the Volapük base and its counterpart in a possible source language. While most entries of the list contain only one possible etymology, there are cases where an accurate connection between the two lexical items cannot be established with confidence, yielding multiple plausible etymologies. For the sake of the research, the pronunciation of the corresponding source items is based on the historical standard pronunciation of these words at the time of the invention of Volapük. In particular, I now use these pronunciation standards for the following languages: English (based on British Received Pronunciation), German (based on Bühnendeutsch by Siebs (1912)), Latin, and Ancient Greek (based on the contemporary realization of Latin and Ancient Greek loanwords in German, indicated by De Jong (1931)).

Furthermore, the gazette *Volapükagased pro Nedänapükans* has been specifically chosen as the source of the wordlist, since it is the only reliable long-term material written in Volapük nulik in the sense that it contains texts of various purposes and themes (e.g., poems, short novels, news, non-fictions, advertisements, etc.) written by *mastans* or authoritative Volapükists, such as *Revidan* 'the reviser of the Volapük language', Arie de Jong, and *Cifal* 'the head of the Volapük movement' at the time, Albert Sleumer.

2.2. DATA ANALYSIS. For ease of analysis, the attached spreadsheet containing the wordlist is imported into Python using Pandas. For the descriptive analysis in Section 4.2, regexes are used to analyze the column of Volapük bases in IPA. All percentages are rounded to the nearest tenth. A Jupyter Notebook "Supplement2" is attached to the thesis in order to provide more detail.

R is used for the analysis of vowel sequences in Section 4.2.4 to calculate cumulative *p* values based on binomial tests, which analyzes whether some vowel sequences are significantly more common than others in Volapük (despite the fact that Volapük has no diphthongs prescriptively, as discussed in Section 3), since some vowel sequences are significantly more common than others. This is accomplished first by calculating the joint probability of having such a vowel sequence (i.e., by multiplying the probabilities of the two independent vowels at their respective position), then computing the cumulative *p* values as well as the adjusted *p* values according to false discovery rate (FDR) adjustment with built-in statistical functions in R. The final *p* values are rounded to the nearest thousandth for the sake of accuracy.

A Word document “Supplement3” containing the commands in R and a spreadsheet “Supplement4” containing the data used for calculating the *p* values are attached to the thesis also to provide more detail.

3. ORTHOGRAPHY. The Volapük alphabet is similar to that of German, as it has adopted the Latin alphabet with three letters with umlauts, ‹ä, ö, ü›. On the other hand, the letters ‹q, w›, though being part of the Latin alphabet, have been left out of Volapük. Therefore, the Volapük alphabet consists of 27 letters in total, 8 of which are vowels, namely ‹a, e, i, o, u, ä, ö, ü›, while the other 19 letters are consonants, namely ‹b, c, d, f, g, h, j, k, l, m, n, p, r, s, t, v, x, y, z›. Most letters are pronounced in the same way as their IPA counterparts are. However, flexibility is also observed to a certain degree in some letters that have alternative ways of pronunciation in practice, as shown in Table 1, Figure 1, and Figure 2.

TABLE 1. The Volapük alphabet.

Letter	IPA representation	Alternative IPA
‹a›	/a/	
‹ä›	/ɛ/	[æ]
‹b›	/b/	
‹c›	/dʒ/	[tʃ]
‹d›	/d/	
‹e›	/e/	
‹f›	/f/	
‹g›	/g/	

‹h›	/h/	
‹i›	/i/	
‹j›	/ʒ/	[ʃ]
‹k›	/k/	
‹l›	/l/	
‹m›	/m/	
‹n›	/n/	
‹o›	/o/	
‹ö›	/ø/	
‹p›	/p/	
‹r›	/r/	[ɹ], [ʁ], [ʀ]
‹s›	/z/	[s]
‹t›	/t/	
‹u›	/u/	
‹ü›	/y/	
‹v›	/v/	
‹x›	/gz/	[ks]
‹y›	/j/	
‹z›	/dz/	[ts]

	Bilabial	**Labial-dental**	**Alveolar**	**Post-alveolar**	**Palatal**	**Velar**	**Uvular**	**Glottal**
Plosive	p b		t d			g k		
Nasal	m		n					
Trill			r				(ʀ)	
Fricative		f v	(s) z	(ʃ) ʒ			(ʁ)	h
Affricate			(ts) dz	(tʃ) dʒ				
Approximant			(ɹ)		j			
Lateral approximant			l					

FIGURE 1. IPA Chart for Volapük consonants.

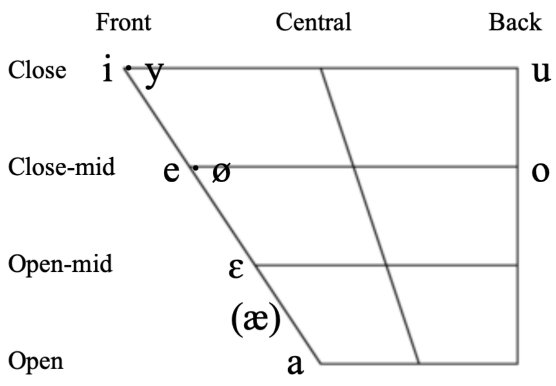

FIGURE 2. IPA Chart for Volapük vowels.

According to De Jong (1931), each letter represents one and only one sound, while every letter must also be pronounced separately from one another prescriptively. This is reflected in the case of vowel sequences such as ‹beit› ‘bite’, in which the vowel sequence ‹ei› is pronounced as two separate vowels /e.i/ (thus /beˈit/) instead of a diphthong. Descriptively, however, this fundamental rule is violated by the letter ‹x›, and to some degree the letter ‹s› as well. The letter ‹x› – represented in IPA as /gz/ – represents a consonant cluster rather than an individual consonant and often occurs in the base form of a word, while the same consonant cluster is represented in some contexts by a combination of letters ‹g› and ‹s›, particularly between the morpheme boundaries, for instance, in the case of plural noun forms (e.g., ‹glügs› /glygz/ ‘churches’, which is formed by the singular from ‹glüg› suffixed by the plural ending ‹s›). The letter ‹s› – prescriptively represented in IPA as /z/ alone – on the other hand, is usually pronounced in practice as [z] whenever possible for the speaker and [s] elsewhere, particularly in a consonant cluster where another voiceless consonant is present (e.g., ‹sag› ‘speech’ would retain its prescriptive way of pronunciation and still be pronounced as [zag], but ‹spad› ‘space’ would be pronounced as [spad] instead of /zpad/, which is hardly pronounceable for most speakers). Another descriptively possible realization of the /z/-phoneme according to Wasilewski (2015) is to substitute it completely with [s], for instance, the letter ‹s› in both ‹sag› and ‹spad› would be realized as [s]; therefore, the two bases here would be pronounced as [sag] and [spad], respectively.

4. STRUCTURAL ANALYSIS OF VOLAPÜK BASES.

4.1. PRESCRIPTIVISM. While the borrowing of Volapük bases has been relatively arbitrary and free, De Jong (1931) nevertheless briefly discusses some prescriptive restrictions on the structure of Volapük bases (mostly nominal ones), which are therefore considered fundamental to the base formation process. These rules are:

1. Bases in general must be short for ease of further derivation.
2. Noun bases must begin and end with consonants for ease of pronunciation in case of further derivation and inflection processes.
3. Uninflected nouns in general (the singular nominative case) cannot end in one of the five sibilants (‹x› included) for ease of pronunciation in case of pluralization, which is realized with the letter ‹s› suffixed to the stem.
4. Although noun bases can begin and, with the exception of the sibilants, end in any consonants, certain monosyllabic letter sequences designated for inflectional or derivational morphemes are expected to be avoided as much as possible during the process of base borrowing in order to avoid

unnecessary ambiguity. For instance, the monomorphemic base ‹famül› ‘family’ should have been theoretically avoided, because it could be otherwise interpreted as the base ‹fam› ‘fame’ suffixed with ‹ül›, a derivational suffix for diminutive with endearment, though the hypothetical complex word *‹fam-ül› does not exist in reality.

4.2. Descriptive Analysis.

4.2.1. Length. The quality of length is examined here in terms of the number of syllables of the bases. The result shown in Table 2 indicates that a significant number of the 1000 bases in data are monosyllabic (68.5%), while another 28.9% of them have two syllables. In addition, the frequency of each syllabic type also decreases as the number of syllables increases. This proves the prescriptive rule 1 in Section 3.1, as proposed by De Jong (1931).

TABLE 2. Frequencies according to the length in syllables.

Type	Subtype	Quantity	Frequency[a]
Monosyllabic	-	685	68.5%
Multisyllabic[b]	-	315	31.5%
	With 2 syllables	*289*	*91.7%*
	With 3 syllables	*25*	*8.0%*
	With 4 syllables	*1*[c]	*0.1%*

[a] The frequency of each subtype is calculated out of the broader type; for instance, 91.7% of the multisyllabic bases have two syllables in them.

[b] 129 (41.0%) of the multisyllabic bases featuring at least one vowel combination as discussed in Section 2.

[c] The only base in data with four syllables is ‹diam-ain› ‘diamond’, which contains the usually inseparable suffix ‹-ain› designated for gems. While it is considered a separate base by Volapükists, it would be more reasonable here to view it as a morphologically complex word consisting of two bound morphemes, namely, ‹diam-› and ‹-ain›, as the former is also observed in words with other possible suffixes such as ‹diam-oin› ‘rough diamond’.

4.2.2. WORD-INITIAL POSITION. A general analysis sheds light upon the possible word-initial elements of any given Volapük bases. The result is shown in Table 3, Table 4, and Table 5.

TABLE 3. Frequencies of word-initial elements.

Element type	Subtype	Quantity	Frequency[a]
Vowel	-	25	2.5%
Consonant	-	975	97.5%
	Single C	*772*	*79.2%*
	CC cluster	*202*	*20.7%*
	CCC cluster	*1*[b]	*0.1%*

[a] The frequency of each subtype is calculated out of the broader type; for instance, 79.2% of the bases starting with at least one consonant begin with only one consonant.

[b] The only attested word-initial cluster with three consonants is ‹str› (IPA: /ztr/), as in ‹stral› ‘radiance’.

TABLE 4. Frequencies of word-initial single consonants.

Consonant in IPA	Quantity	Frequency (out of 772)
/z/	83	10.8%

/l/	72	9.3%
/d/	69	8.9%
/m/	68	8.8%
/k/	60	7.8%
/v/	54	7.0%
/f/	52	6.7%
/t/	47	6.1%
/n/	44	5.7%
/b/	43	5.6%
/p/	40	5.2%
/g/	31	4.0%
/r/	28	3.6%
/ʒ/	20	2.6%
/dʒ/ (affricate ‹c›)	18	2.2%
/dz/ (affricate ‹z›)	18	2.3%
/j/	16	2.1%
/h/	9	1.2%

TABLE 5. Frequencies of word-initial CC clusters.

CC cluster in IPA	Quantity	Frequency (out of 202)
/zt/	32	15.8%
/pl/	19	9.4%
/pr/	19	9.4%
/kl/	16	7.9%
/fl/	15	7.4%
/bl/	14	6.9%
/gl/	11	5.5%
/gr/	10	5.0%
/kr/	8	4.0%

/zp/	8	4.0%
/dr/	7	3.5%
/fr/	7	3.5%
/zk/	7	3.5%
/zl/	6	3.0%
/zv/	6	3.0%
/tr/	6	3.0%
/br/	5	2.5%
/zm/	3	1.5%
/gz/ (written as ‹x›)	2	1.0%
/kv/	1	0.5%

Among all 975 bases beginning with at least one consonant, the analysis reveals a wide variety of grammatically possible consonant clusters at word-initial position, including:

- 20 clusters consisting of two consonants: /br/, /bl/, /dr/, /fl/, /fr/, /gl/, /gr/, /gz/, /kl/, /kr/, /kv/, /pl/, /pr/, /tr/, /zk/, /zl/, /zm/, /zp/, /zt/, /zv/
- 1 cluster consisting of three consonants: /ztr/.

Most of these clusters follow the Sonority Sequencing Principle (i.e., increasing sonority in the onset part of a syllable, the closer the consonant is to the nucleus, as discussed in Section 4.2.2) and thus well observed and predictable in natural languages as well. The only exceptions to this principle are those clusters beginning with /z/ and then followed by a voiceless stop (/zk/, /zp/, and /zt/), which are also particularly interesting as they feature individually a voiced-voiceless pair, which would undergo assimilation in most natural languages. While this phenomenon is rarely seen outside Volapük, it could potentially be attributed to prescriptivism, as proposed in Section 2. In other words, while the letter ‹s› is prescriptively /z/, it is often pronounced as [s] in reality. Thus, these three clusters (and to certain degree also /zl/, /zm/, /zv/) would be comparable with the s-clusters (/sk/, /sp/, /st/) that are more often encountered in natural languages.

This nonprescriptive [s] helps to explain the apparent violation of Sonority Sequencing Principle, because [s] has been described as a common onset adjunct, to which the principle does not apply (Gierut 1999, Pan & Snyder 2004, Selkirk 1982). Therefore, it is possible to categorize all possible word-initial clusters in Volapük into three groups: those obeying the Sonority Sequencing Principle and without the initial /z/ (/br/, /bl/, /dr/, etc.), those disobeying the principle (/zk/, /zp/, /zt/, /ztr/), and those obeying the principle and with the initial /z/ (zl/, /zm/, /zv/). In the case of the first group, all such clusters can be regarded as the onset of the word-initial syllable, as they fully conform with the principle, while in the second group, the word-initial /z/ would be more reasonably regarded as a syllable adjunct that does not form part of the onset of the word-initial syllable, and the principle would therefore still be applicable to the actual onset (in this case only /k/, /p/, /t/, and /tr/). The last group, however, may be somewhat ambiguous, as technically these clusters could still confirm with the principle (nasal /m/ and liquid /l/ have more sonority than fricative /z/, while /v/ is also a fricative as /z/). Nonetheless, the data shows that no other fricatives except /z/ can be followed by a nasal or another fricative, and it is not possible to have a fricative-nasal-liquid sequence in Volapük, either. Thus, it is probably more reasonable to presume that /z/ is also a syllable adjunct in this group, as its appearance in these clusters seems so unique that it cannot be replaced by any other similar elements that have similar sonority (i.e., with the same distance).

Other consonant clusters observed at word-initial position that are rare in natural languages include, most notably, /gz/ and /kv/. The first cluster is an IPA realization of the Volapük letter ‹x›, as discussed in Section 2. The only examples on the list include the words ‹xam› 'examination' and ‹xän› 'ascent', both of which are ultimately borrowed from Latin, where the counterparts of this /gz/ sequence, namely the /ks/ and the /sk/ sequences, appear in word-middle position. In other words, the vowel apheresis and consonant metathesis in Volapük word-borrowing process enables the consonant cluster /gz/ to stand at word-initial position, while its counterparts in the origin language, namely Latin, is expected to be observed in a different environment (i.e., at word-internal position). The second cluster /kv/ is featured only in the nominal base ‹kvär› 'acorn'. Ultimately borrowed from Latin, the Volapük sequence /kv/ is the counterpart of the Latinate cluster ‹qu› (in this case from ‹quercus› 'oak'), which is realized as /kv/ according to German convention (thus /ˈkverkus/). Compared to other possible clusters, these two only comprise a very small portion (0.990% and 0.495% out of all bases beginning with a cluster of two consonants, respectively), thus indicating their rarity in the Volapük language.

According to the prescriptive rule 2 in 2.1, nominal bases should never begin with a vowel, yet 25 out of the 1000 bases in the data seem to bear this characteristic. Therefore, a further analysis on these 25 bases is conducted in order to determine their respective part of speech. The result given in Table 6 proves the same prescriptive rule, as none of these 25 bases that begin with a vowel is nominal.

TABLE 6. Types and frequencies of the bases beginning with a vowel.

Part of speech	Quantity	Frequency (out of 25)
Conjunction	7	28.0%
Preposition	7	28.0%
Adverb	5	20.0%
Pronoun	4	16.0%
Interjection	1	4.0%
Definite article ‹el›	1	4.0%

4.2.3. WORD-FINAL POSITION. Possible elements at word-final position are analyzed in the same way as those at word-initial position. The result of the analysis is shown in Table 7, Table 8, and Table 9.

TABLE 7. Frequencies of word-final elements.

Element type	Subtype	Quantity	Frequency[a]
Vowel	-	80	8.0%
Consonant	-	920	92.0%
	Single C	894	97.2%
	CC cluster	26	2.8%

[a] The frequency of each subtype is calculated out of the broader type; for instance, 97.2% of the bases ending in at least one consonant end in only one consonant.

TABLE 8. Frequencies of word-final single consonants.

Consonant in IPA	Quantity	Frequency (out of 894)
/d/	151	16.9%
/t/	128	14.3%
/n/	125	14.0%
/l/	110	12.3%
/m/	80	9.0%
/f/	61	6.8%
/g/	53	5.9%
/r/	44	4.9%
/k/	38	4.3%
/b/	37	4.1%
/p/	32	3.6%
/v/	24	2.7%
/z/[a]	11	1.2%

[a] /z/ is the only possible word-final sibilant according to the data.

TABLE 9. Frequencies of word-final CC clusters.

CC cluster in IPA	Quantity	Frequency (out of 26)
/zt/	5	19.2%
/ld/	3	11.5%
/rn/	3	11.5%
/rm/	2	7.7%
/nd/	2	7.7%
/zk/	2	7.7%
/nk/	1	3.9%
/rt/	1	3.9%
/lt/	1	3.9%
/mp/	1	3.9%
/nt/	1	3.9%

/rk/	1	3.9%
/kt/	1	3.9%
/rg/	1	3.9%
/lf/	1	3.9%

Thus, the majority (89.4%) of the possible elements at word-final position, as revealed in data, include all 12 non-sibilants, while another 10.6% of the data are occupied by the vowels, the sibilant /z/, as well as 15 consonant clusters /kt/, /ld/, /lf/, /lt/, /mp/, /nd/, /nk/, /nt/, /rg/, /rk/, /rm/, /rn/, /rt/, /zk/, and /zt/. Compared to the wide variety at word-initial position, it seems that at word-final position, consonant clusters only comprise a significantly smaller portion of the data (20.3% word initially compared to only 2.8% word finally) and have less variety (21 types at word-initial compared to only 15 at word-final), though most of these clusters at word-final position (except /kt/, which is a plateau) conform with the Sonority Sequencing Principle (in this case decreasing sonority from the nucleus to the end of the coda) and are thus typical in natural languages. Furthermore, none of these clusters in word-final position seems to stand out in terms of its frequency: namely, their occurrences are only sporadic and scattered (while the most frequent cluster at word-initial position, /zt/, is found in 3.2% of the data, the most frequent one at word-final position, also /zt/, is only detected in 0.5% of all data). Therefore, it is reasonable to conclude that Volapük, a constructed language, features the same concept of prominent position as natural languages in the sense that it favors consonant clusters more at word-initial position than at word-final position. In other words, natural languages are found to contain a greater variety of segment types at prominent prosodic positions than non-prominent positions (De Lacy 2000), and the same is observed in Volapük word-initial syllables as opposed to word-final ones.

It remains in question, however, whether such a phenomenon is attributed to the a posteriori properties of Volapük, namely, unintentionally influenced by the origin languages, or if it is deliberately adopted and maneuvered during word-borrowing process. Further research must be conducted between Volapük bases and their counterparts in respective languages of origin – if the data in origin languages already show this tendency, then probably it is only coincidental; if it is only featured in Volapük data, namely, the origin of these bases do not or rarely have this tendency, then the property must have been deliberately added in the word-borrowing process.

A similar analysis on part of speech is conducted on those bases ending in a vowel and in /z/, which is the only word-final sibilant present in data. This analysis aims at prescriptive rules 2 and 3 described in Section 3.1, which propose that nominal bases should end in neither a vowel nor a sibilant. The result is given in Table 10 and Table 11, according to which the prescriptive rules 2 and 3 also seem to hold true within the data, as none of these 91 bases ending in a vowel or /z/, the only possible sibilant according to data, is nominal, namely, able to be pluralized by suffixing another /z/ at the word-final position.

TABLE 10. Types and frequencies of the bases ending in a vowel.

Part of speech	Quantity	Frequency (out of 80)

Preposition	39	48.8%
Adverb	25	31.3%
Conjunction	11	13.8%
Interjection	5	6.3%

TABLE 11. Types and frequencies of the bases ending in /z/.

Part of speech	Quantity	Frequency (out of 11)
Preposition	7	81.8%
Adverb	3	27.3%
Conjunction	1	9.1%

4.2.4. VOWELS AND VOWEL SEQUENCES. As discussed in Section 2, all eight vowels are potentially able to serve as the nucleus in any given syllable in Volapük. The frequency of each individual vowel out of the sample of 1000 bases (yielding 1342 syllables in total) is given in Table 12. According to the result, /e/, /i/, and /a/ appear the most frequently among all vowels, while the three umlaut-having letters (/ɛ/, /y/, and /ø/) appear less frequent than the other five vowels in general.

TABLE 12. Frequencies of vowels.

Vowel in IPA	Quantity	Frequency (out of 1342)
/e/	273	20.3%
/i/	272	20.3%
/a/	227	16.9%
/o/	183	13.6%
/u/	139	10.4%
/ɛ/	96	7.2%

/y/	82	6.1%
/ø/	70	5.2%

While a sequence of two or more vowels is prescriptively pronounced separately as discussed in Section 2 according to De Jong (1931), further analysis on vowel sequences shown in Table 13 indicates the possibility that the selection of certain sequences during the word-borrowing process from natural languages may be structural and systematic. Among all possible vowel sequences, it is observed that 3 of them, namely /ei/, /ie/, and /ai/, appear significantly more often (i.e., has an FDR-adjusted *p* value smaller than or equal to 0.05) than they would appear if randomly selected and combined into this sequence from the 1342 vowel tokens. The etymologies of the bases where these vowel sequences are present further suggest that many bases exhibiting these sequences realize them as the diphthong /aɪ/ in their respective source language, for instance, ‹beit› ‘bite’, ‹viet› ‘white’, ‹bai› ‘by’. Thus, despite the fact that they are prescriptively realized and pronounced as two separate vowels, this systematic borrowing of certain vowel sequences might, conversely, indicate an immature stage of diphthongization in the Volapük language.

TABLE 13. Frequencies and *p* values of vowel sequences.

Vowel sequence in IPA	Quantity	Frequency (out of 130)	Unadjusted *p* value	Adjusted *p* value
/ei/	28	21.5%	<0.001	<0.001
/ie/	17	13.1%	<0.001	<0.001
/ai/	11	8.5%	0.005	0.047

/ia/	9	6.9%	0.034	0.238
/ea/	7	5.4%	0.155	0.494
/oe/	6	4.6%	0.159	0.494
/io/	6	4.6%	0.159	0.494
/oa/	5	3.9%	0.181	0.507
/au/	5	3.9%	0.087	0.485
/ui/	4	3.1%	0.292	0.667
/iɛ/	4	3.1%	0.133	0.494
/eo/	3	2.3%	0.708	0.964
/ɛi/	3	2.3%	0.310	0.667
/ya/	2	1.5%	0.374	0.747
/uɛ/	2	1.5%	0.231	0.588
/ey/	2	1.5%	0.463	0.789
/ue/	2	1.5%	0.760	0.964
/eu/	2	1.5%	0.760	0.964
/oi/	2	1.5%	0.882	0.987
/ou/	2	1.5%	0.545	0.847
/ye/	1	0.8%	0.792	0.964
/ae/	1	0.8%	0.989	0.989
/ao/	1	0.8%	0.951	0.987
/iu/	1	0.8%	0.937	0.987
/ua/	1	0.8%	0.892	0.987
/yɛ/	1	0.8%	0.406	0.758
/iø/	1	0.8%	0.763	0.964
/ɛɛ/	1	0.8%	0.479	0.789

4.2.5. SUMMARY OF FINDINGS. Based on my descriptive analysis, I conclude that most of the prescriptive rules in De Jong (1931) are being well observed by the actual data, in particular the prescriptive rules 2 and 3 mentioned in Section 4.1, to which the data yields no exceptions. However, there also seems to be findings that are not explicitly mentioned in Volapük grammar, most notably the significantly low frequency of complex coda at word-final position (2.8% as described in Section 4.2.3), while prescriptivism only requires a mandatory non-sibilant word-final coda. Furthermore, Section 4.2.4 on vowel sequences also seems to suggest a de facto violation of prescriptivism, as the descriptive analysis indicates a possible immature stage of diphthongization in Volapük, while diphthongs are prescriptively forbidden in Volapük based on Section 4.1 and De Jong (1931).

5. SYLLABLE TEMPLATE. Syllable templates are proposed in Selkirk (1982) as a method to illustrate the mandatory syllable structure in a given language. Based on the data from the wordlist, it is therefore possible to construct the generalized syllable template for Volapük as shown in Figure 3, which is applicable to the majority of the data in this analysis:

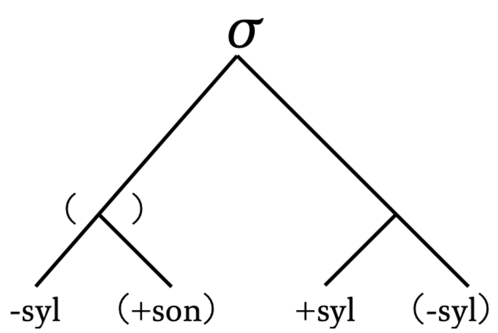

FIGURE 3. Generalized syllable template for Volapük.

This template concludes that, in general, a Volapük syllable consists of only one mandatory element, namely nucleus, whereas onset (consisting of either one consonant, or optionally a consonant cluster in which the second consonant is a sonorant) and coda are both optional to the syllable structure. For instance, true Volapük bases such as /lam.pør/ 'emperor' and /lo.to.graf/ 'orthography' will fit into this template, while hypothetical syllables whose onset is comprised of two obstruents such as */ptaf/, or those whose coda is comprised of more than one consonant such as */tapf/, would not.

However, based on the analysis in Section 4.2, this generalized syllable structure requires modifications at word boundaries for nominal bases, as additional restrictions are mandatory at the word-level. This is illustrated through the exemplar monosyllabic template in Figure 4, in which the previously optional onset becomes mandatory now at word-initial position, meanwhile coda consisting of a non-sibilant also becomes required at word-final position. For instance, this explains Volapük bases such as /plad/ 'place' and /zek/ 'consequence', meanwhile not acknowledging the hypothetical ones whose only syllable consists of either no onset (e.g., */ap/) or a complex or a sibilant coda (e.g., */tapf/ or */taz/). It needs to be noted that this modified template applies not only to monosyllabic bases but also to multisyllabic ones at the word-level; the figure features the former only for convenience's sake.

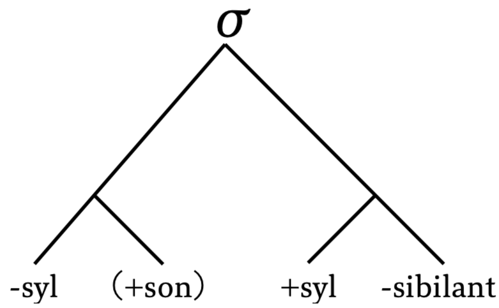

FIGURE 4. Syllable template for Volapük at the word-level.

While both templates account for a majority of the Volapük bases in data, they also encounter several exceptions, including the nominal bases whose coda consists of a consonant cluster instead of a single consonant (which occupy 2.8% of the data, as analyzed in Table 7). Additionally, the templates suggest that Volapük syllables fully conform with Sonority Sequencing Principle, a conclusion which the word-initial /z/-obstruent (/zk/, /zp/, /zt/) clusters seem to violate, as analyzed in Section 4.2.2. This is justified by Selkirk (1982), who proposes that an additional auxiliary template is adopted for the structurally comparable English onset cluster comprised of /s/ plus obstruent so that the word-initial /s/ plus the additional obstruent together form a single obstruent instead of a consonant cluster, as shown in Figure 5 (reproduced from Selkirk 1982).

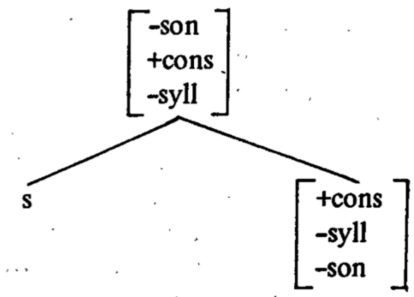

FIGURE 5. Auxiliary template for s-clusters (reproduced from Selkirk 1982).

Thus, the /s/-cluster as a single entity only occupies one slot of the template (i.e., the first consonant) rather than two separate slots, enabling the application of the Sonority Sequencing Principle as the sonority based on the template now is increasing from the onset slots of a syllable towards the nucleus. Similarly, the violation of Sonority Sequencing Principle by word-initial /z/-clusters in Volapük (descriptively realized as [s] in this case, as analyzed in Section 3) can also be justified in the same way, illustrated in Figure 6 through the representation of the base ‹stral› ‘radiance’.

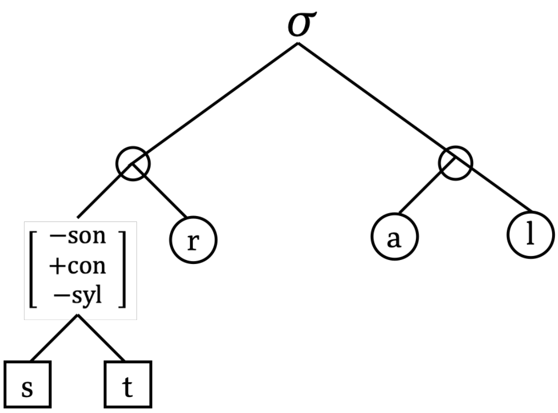

FIGURE 6. Representation of the base ‹stral› under syllable templates.

6. Loanword Adaptation in Volapük.

6.1. Loanword Strategy. The third part of the thesis combines the methods of loanword adaptation model and Optimality Theory (OT, Prince and Smolensky 1993) in order to analyze how Volapük nominal bases are derived from their counterparts in natural languages. The OT results indicate a great degree of consistency with the descriptive analysis in Section 4, both of which exhibit the preference of Volapük for a certain syllable shape, as generalized by the syllable templates in Section 5.

In the case of natural languages, loanwords refer to those words that are borrowed from other natural languages; therefore, in order to conform with the phonotactics of the target language, the words in source languages have to undergo certain prosodic and segmental changes, namely the "adaptation" process. The model in Silverman (1992) assumes that this adaptation includes first a perceptual and then a phonological phase: foreign input from the native speaker of the source language, namely, the pronounced word, is first perceived by a native speaker of the target language, then serving as the underlying form that undergoes phonological changes, the output of which is the actual loanword appearing in the target language, namely, the surface form of the phonological process.

This adaptation model for loanwords in natural languages is somewhat relatable to the general word-borrowing process in constructed languages, as a posteriori constructed languages, including Volapük, borrow most of their lexical bases from existing natural languages. However, it has to be noticed that the word-borrowing process in constructed languages is inherently different from that in natural languages, in the sense that it is a much more conscious process and therefore lacks the perceptual phase of the adaptation process. The whole word-borrowing process seems more arbitrary at least from the surface, since the target bases of constructed languages are consciously modified and borrowed from existing natural languages by their creators, whereas the adaptation process in natural languages are much more emergent, that is, involving less conscious intervention. Furthermore, while the pronounced word in a target language still exists, it is not "perceived" first by any native speakers of the target constructed language, as the language cannot have any native speakers before the lexicon of the language, which is composed exclusively of such "loanwords," is completed. Furthermore, even if the construction is completed, only very few constructed languages, for instance, Esperanto, have native speakers at all, while they are often found to have developed their own set of particular synchronic grammar, thus understanding and perceiving the language in a substantially different way than those who learn the language as a second language, including the creator of the constructed language (i.e., those who construct the lexicon of the language from their knowledge of natural languages) as well (Bergen 2001). While Volapük does not have a record of any native speakers throughout its history, a similar phenomenon (i.e., a particular set of native synchronic grammar that is different from that of L2 speakers) is likely to be observed among hypothetical Volapük native speakers as among native Esperanto speakers due to their similarity as a posteriori constructed languages.

Therefore, in the case of any constructed languages, including Volapük, one might hypothesize the existence of a consistent grammatical process of curation from source language to target language, which is analogous to the loanword adaptation model in natural languages, provided that that the underlying form for Volapük bases are the pronounced form in the source language. If a coherent OT grammar can be established, then this consistent process is also proven to exist.

6.2. CONSTRAINTS. A series of constraints needs to be established for this analysis, which examines the data using the methods of Optimality Theory. As Prince and Smolensky (1993/2004) and Yip (1993) suggest, languages strive to achieve well-formedness through the filter of constraints. Unlike the syllable templates in Section 5, these constraints are typically violable, and it is often necessary to establish a set of rankings for them, in which the higher ranked constraints should be prioritized and more likely to be observed than the lower ranked ones.

Universally, the set of constraints can be categorized into two groups, namely faithfulness and markedness constraints. For this particular analysis, it is proposed that faithfulness constraints contain two different subcategories, namely Max (prohibiting deletion) and Dep (prohibiting epenthesis), while theoretically segmental constraints (i.e., concerning change in specific phonemes) are present as well. Markedness constraints, on the other hand, are referred to mostly as prosodic constraints (i.e., concerning the generalized syllable structure) in this section.

6.2.1. FAITHFULNESS CONSTRAINTS. One particular type of constraints that are universally featured in an OT analysis is the faithfulness constraints as proposed by Prince and Smolensky (1993/2004). As McCarthy (2008) explains, these constraints "prohibit differences between input and output," and in the case of the adaptation process in Volapük, they are utilized to examine the conformity between the origin word in the source language, namely the input, and the target word in Volapük, namely the output. In this analysis, two types of faithfulness constraints are featured, namely Max and Dep, both proposed by McCarthy (2008). They are then further classified into MaxC, MaxV, DepC, and DepV as follows:

1. MaxC: Assign a violation for every consonant in input that is being deleted in output.
2. MaxV: Assign a violation for every vowel in input that is being deleted in output.
3. DepC: Assign a violation for every consonant in output that is not featured in input.
4. DepV: Assign a violation for every vowel in output that is not featured in input.

6.2.2. PROSODIC CONSTRAINTS. Prosodic constraints regulate the syllable structure to which the Volapük bases are expected to conform during the adaptation process. Based on the syllable templates in Section 5, the following five prosodic constraints are proposed to be relevant to this OT analysis:

1. Onset: Assign a violation for every word-initial vowel.
2. WordCoda: Assign a violation for every word-final vowel or sibilant. This corresponds to the coda part of the word-level syllable template shown in Figure 4 of Section 5.
3. NoCoda: Assign a violation for every syllable with a coda.

4. *CompCoda: Assign a violation for every complex coda, that is, consisting of a consonant cluster.
5. SSP: Assign a violation for every consonant cluster that violates Sonority Sequencing Principle. The sonority sequence in Figure 7 is proposed for Volapük phonemes in general, and a violation is assigned for every cluster that exhibits a reversed sonority sequence or a plateau contrary to the figure.

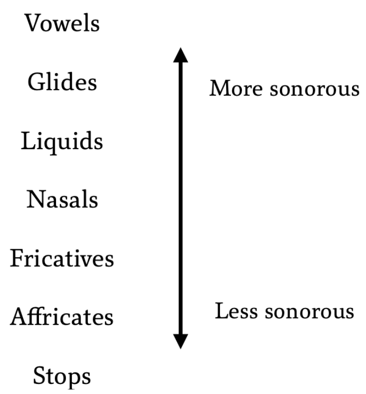

FIGURE 7. Sonority sequence for Volapük phonemes.

6.2.3. SEGMENTAL CONSTRAINTS. Unlike Max/Dep and the prosodic constraints whose ranking information is obeyed by most Volapük bases, the hypothesized segmental constraints seem to be more volatile in the sense that their ranking information is expected to feature much more contradiction. This is due to the rather arbitrary and, in many cases, completely random occurrence of segmental changes during the actual adaptation process. It is therefore difficult to establish any specific segmental constraints as the preference over these segmental changes seems to be at least obscure, sometimes even contradictory according to the data. Some of the frequently observed segmental changes during the Volapük base adaptation process include:

1. Voicing and devoicing at word-initial and final positions
2. Avoidance of ‹r› and ‹h›
3. Change in vowel quality
4. Consonant epenthesis at word-final position

These changes do not exhibit any regular pattern in the sense that the preference for their occurrence is mostly due to extralinguistic motives that are often irrelevant in the case of natural languages, while pure arbitrariness sometimes also contributes to this irregularity. Within the frame of the data, the following possibilities attempt to account for these changes:

1. Avoiding homonyms: According to De Jong (1931), Volapük does not allow the existence of homonyms. Therefore, origin words must undergo additional phonological changes if their target counterpart resembles any existing Volapük bases. This is likely to be the case with the voicing in [bleg] ‘the quality of being bent’ (from Latin /plekt-/) vs. [plek] ‘praying’, and the vocalic change in [klyl] ‘clearness’ (from English /klɪə/) vs. [klil] ‘lightness’.
2. Conforming to orthography: According to De Jong (1931), Volapük bases may rather adhere to the written form of their origin counterpart, but not the spoken form. This is likely to be the case with the word-final voicing in [veg] ‘road’ from German ‹Weg› /veːk/, the monophthongization in [lif] ‘life’ from English ‹life› /laɪf/, and the /t/-epenthesis in [gyt] ‘sense’ from French ‹goût›

/gu/.

3. Disfavor of /r/ and /h/ in Volapük rigik: According to the etymological data on Volapük rigik from the word list, the etymology of many frequently used Volapük nulik bases actually traces back to Schleyer's version of the language, namely Volapük rigik, which was mainly used in late 1880s. Schleyer (1880) notes that the frequencies of the letter ‹r› and ‹h› are intentionally reduced for the ease of pronunciation for East Asians, while many of the Volapük rigik bases that underwent such changes had been so frequently used in ordinary speech that De Jong, the creator of Volapük nulik, did not further modify them during his revision of the language in early 1930s. This is the case with [flen] 'friend' from English /fɹɛnd/, and [lad] 'heart' from English /hɑːt/.
4. Pure arbitrariness: For many other Volapük bases, the phonological changes they have undergone do not seem to be reasonable at first glance. They could be accredited to the preference of the creators or to the arbitrary nature of Volapük as a constructed language. This is the case with the word-final devoicing in [gløp] 'ball' from English /gləʊb/, the word-initial devoicing in [fɛlid] 'valley' from English /væli/, the /p/-epenthesis in [bap] 'inferiority' from French /bɑ/ (‹bas›), and the vocalic change in [gyt] 'sense' from French /gu/ (‹goût›).

6.3. CONSTRAINT TABLEAUX. The exemplar Volapük nominal bases in Table

14 are chosen for the constraint tableau, which includes their individual characteristics (in particular those concerning the syllable structure), the source words in IPA, as well as the English translations. Stress is ignored as it is irrelevant to this analysis.

TABLE 14. Exemplar Volapük nominal bases.

Target word	Source word	English translation	Characteristics
[bel]	German: /bɛrk/	mountain	Coda cluster simplification via deletion
[fɛlid]	English: /væli/	valley	Word coda epenthesis
[leig]	English: /əlaɪk/	sameness	Vowel deletion
[løp]	English: /ʌp/	upper side	Onset epenthesis
[spad][a]	French: /ɛspas/	space	Word coda modification (s→d) Vowel deletion

[a] prescriptively /z/.

Given the exemplar bases in Table 14, the constraint tableaux shown in (1)-(5) are constructed, which gives information on specific rankings of various constraints, in particular those between the faithfulness and prosodic constraints. DepV, however, is only featured in few of the words, and does not seem to be representative. Segmental constraints (i.e., indicated by specific phonemic changes), on the other hand, are not present in this tableau due to their complete arbitrariness, as discussed in Section 6.2.3.

(1) Constraint tableau for /bɛrk/ ~ [bel] 'mountain':

bɛrk 'mountain'	WordCoda	Onset	*CompCoda	NoCoda	DepC	MaxV	SSP	MaxC	DepV
→ bel				*				*	
belk			W*	*				L	

In the case of /bɛrk/, the candidate [bel] with a simplified coda is preferred over the faithful candidate *[belk], whose coda contains a consonant cluster. Therefore, *CompCoda is ranked higher than MaxC.

(2) Constraint tableau for /væli/ ~ [fɛlid] 'valley':

væli 'valley'	WordCoda	Onset	*CompCoda	NoCoda	MaxV	DepC	SSP	MaxC	DepV
→ fɛlid				*		*			
fɛli	W*			L		L			
fɛl				*	W*	L			

In the case of /væli/, the candidate [fɛlid] with an epenthesized coda is preferred over the faithful candidate *[fɛli], which lacks a word-final coda, as well as the candidate *[fɛl], which acquires its coda through final vowel deletion. Therefore, WordCoda is ranked higher than DepC and NoCoda, and MaxV is ranked higher than DepC.

(3) Constraint tableau for /əlaɪk/ ~ [leig] 'sameness':

əlaɪk 'sameness'	WordCoda	Onset	*CompCoda	NoCoda	DepC	MaxV	SSP	MaxC	DepV
→ leig				*		*			
aleig		W*		*		L			

In the case of /əlaɪk/, the candidate [leig], which acquires its word-initial onset through vowel deletion, is preferred over the faithful candidate *[aleig], which begins with a vowel. Therefore, Onset is ranked higher than MaxV.

(4) Constraint tableau for /ʌp/ ~ [løp] 'upper side':

ʌp 'upper side'	WordCoda	Onset	*CompCoda	NoCoda	DepC	MaxV	SSP	MaxC	DepV
→ løp				*	*				
øp		W*		*	L				

In the case of /ʌp/, the candidate [løp], which acquires its onset through consonant epenthesis, is preferred over the faithful candidate *[øp], which begins with a vowel. Therefore, Onset is ranked higher than DepC.

(5) Constraint tableau for /ɛspas/ ~ [spad] 'space':

ɛspas 'space'	WordCoda	Onset	*CompCoda	NoCoda	DepC	MaxV	SSP	MaxC	DepV
→ spad				*	*	*	*	*	
spas	W*			*	L	*	*	L	
es.pas	W*	W*		W**	L	L	L	L	
e.spas	W*	W*		*	L	L	*	L	

les.pad				W**	W**	L	L	*	
le.spad				*	W**	L	*	*	

In the case of /ɛspas/, the winning candidate [spad] undergoes a twofold modification from its input. It acquires the onset through word-initial vowel deletion, while the coda is also modified as the original sibilant /s/ is deleted and an additional /d/ is epenthesized to the end (this would otherwise be hypothesized that the final /s/ is realized as [d] via a violation of Ident, but this alternate assumption contains too much arbitrariness to be modeled due to segmental changes discussed in Section 6.2.3). The losing candidates, on the other hand, either maintain the faithfulness, that is, possessing the word-initial vowel or the word-final sibilant, or acquire the word-initial onset via /l/-epenthesis. While the winning candidate does not obey Sonority Sequencing Principle, the losing candidate may obey it depending on its specific syllable division. Therefore, the general trend exhibited from this constraint tableau suggests that WordCoda and Onset in general are ranked higher than SSP and the faithfulness constraints, in which DepC is ranked higher than MaxV due to the preference of word-initial vowel deletion over consonant epenthesis.

6.4. RANKING DISCUSSION. From the constraint tableaux in Section 6.3, the individual ranking information as shown in Table 15 can be summarized:

TABLE 15. Ranking information.

Exemplar base	Extracted ranking information
[bel]	*CompCoda > MaxC

[fɛlid]	WordCoda > DepC
	WordCoda > NoCoda
	MaxV > DepC [a]
[leig]	Onset > MaxV
[løp]	Onset > DepC
[spad]	WordCoda > DepC
	WordCoda > MaxC
	DepC > MaxV [a]
	Either WordCoda > DepC, MaxC, MaxV, SSP; or Onset > DepC, MaxC, MaxV, SSP; or NoCoda > DepC, MaxC, MaxV, SSP
	Either WordCoda > DepC, MaxV, MaxC; or Onset > DepC, MaxV, MaxC
	Either NoCoda > MaxV, SSP; or DepC > MaxV, SSP

[a] Contradictory ranking information.

Therefore, the constraints illustrated in this OT analysis seem to indicate a trend in rankings so that the prosodic ones, in particular Onset, *CompCoda, and WordCoda, are ranked the highest among all constraint types, followed by NoCoda, whereas the faithfulness constrains as well as SSP are ranked lower in general. This corresponds to the prescriptive rules presented in Section 4.1, suggesting that all nominal bases in Volapük must begin with a consonant and end in a non-sibilant, and is further supported by the descriptive data analysis on word-initial and word-final positions presented in Section 4.2.2 and 4.2.3, respectively, as the majority of the data possess a word-initial onset and a simple word-final coda that is not a sibilant, which is made mandatory for all monosyllabic nominal bases in syllable templates proposed in Section 5.

Nonetheless, there seems to be an irreconcilable contradiction in this ranking table between DepC and MaxV. In the case of /væli/ ~ [fɛlid], MaxV is ranked higher than DepC, as the winning candidate acquires its word-final coda through consonant epenthesis but not vowel deletion. However, in the case of /ɛspas/ ~ [spad], DepC is ranked higher than MaxV, as the winning candidate acquires its word-initial onset through vowel deletion but not consonant epenthesis. This contradiction is examined by Recursive Constraint Demotion (RCD) as proposed in Prince (2009), the result of which is illustrated in Figure 8 and further confirms the contradiction.

			Stratum I			STR II	Stratum III		Stratum IV		
Input	Winner	Loser	WordCoda	Onset	*CompCoda	NoCoda	DepC	MaxV	SSP	MaxC	DepV
væli	fɛlid	fɛli	W			L	L				
ɛspas	spad	spas	W				L			L	
ɛspas	spad	es.pas	W	W		W	L	L	L	L	
ɛspas	spad	e.spas	W	W			L	L		L	
əlaɪk	leig	aleig		W				L			
ʌp	løp	øp		W							
bɛrk	bel	belk			W					L	
ɛspas	spad	les.pad				W	W	L	L		
ɛspas	spad	le.spad					W	L			
væli	fɛlid	fɛl					L	W			

FIGURE 8. Constraint tableaux examined by Recursive Constraint Demotion (RCD).

While it is possible to attribute this contradiction simply to the artificial nature of Volapük, namely, to the pure arbitrariness due to the manifold modifications on the lexicon of the language by more than one Volapükists during a 50-year period of time, a more logical explanation could also be lying behind this contradiction. As discussed in Section 6.2.3, a unique feature of the word-borrowing process in Volapük that cannot be illustrated by OT analysis is the mandatory prevention of homonyms, namely, a lexical item is expected to have one and only one semantic meaning for the sake of clearness. While this prescriptive rule is often achieved by segmental modification, it also seems to be observed in the case of /væli/ ~ [fɛlid], as the losing candidate [fɛl] 'the act of felling (a tree)' (as well as potentially [vɛl] 'selection' if the impact of the segmental change between /v/~[f] is taken into consideration) is already an existing base in Volapük. As a result, the only candidate that both resembles the input and prevents the occurrence of a homonym, namely, [fɛlid], must be preferred over any other candidates. A unique feature of Volapük as a constructed language, this cannot simply be illustrated in OT as an individual constraint, nor is it featured so explicitly and mandatorily in any natural languages.

7. CONCLUSION. As an a posteriori constructed language, Volapük serves as an intriguing counterpart to the common natural languages, which is reflected by its word-borrowing process. On one hand, this process can be explained by analogy with loanword adaptation and further analyzed with Optimality Theory, similar to any natural languages; on the other hand, the result from the analysis indicates a unique arbitrariness absent from natural languages. Unlike natural languages, the prescriptive rules of Volapük always hold true for every lexical item in data, and virtually no exception to these rules is observed, yet prescriptivism is preferred to such a degree that in many cases, arbitrariness might actually play an important role so that other characteristics or constraints expected in natural languages could be ignored.

References

BERGEN, BENJAMIN K. 2001. Nativization processes in L1 Esperanto. *Journal of Child Language* 28(3).575–95.

DE JONG, ARIE. 1931. *Gramat Volapüka*. Leiden: E. J. Brill.

DE LACY, PAUL. 2000. Markedness in prominent positions. *MIT Working Papers in Linguistics* 53–66.

GIERUT, JUDITH A. 1999. Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research* 42(3).708–26.

MCCARTHY, JOHN J. 2008. *Doing Optimality Theory: Applying theory to data*. Oxford: Blackwell.

PAN, NING, and WILLIAM SNYDER. 2004. Acquisition of /s/-initial clusters: A parametric approach. *Proceedings of the 28th Annual Boston University Conference on Language Development* 436–46.

PRINCE, ALAN. 2009. RCD: The movie; ROA-1057. Online: http://roa.rutgers.edu/article/view/1087

PRINCE, ALAN, and PAUL SMOLENSKY. 1993. Optimality Theory: Constraint interaction in generative grammar; RuCCS-TR-2; CU-CS-696-93. *Technical Report 2*. Online: http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF

PRINCE, ALAN, and PAUL SMOLENSKY. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

SCHLEYER, JOHANN M. 1880. *Volapük die Weltsprache: Entwurf einer*

Universalsprache für alle Gebildete der ganzen Erde. Sigmaringen: C. Tappen.

SELKIRK, ELISABETH A. 1982. The syllable. *The Structure of Phonological Representations* 337–83.

SILVERMAN, DANIEL. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9(2).289–328.

SIEBS, THEODOR. 1912. *Deutsche Bühnenaussprache* (10th ed.). Bonn: Albert Ahn.

WASILEWSKI, IGOR. 2015. *Volapük paspiköl* [Video file]. In *Volapükanef*. Online: http://volapuk.temerov.org/Volapükanef/vol/mödamedäd.php

YIP, MOIRA. 1993. Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2(3).261–91.

ZHANG, YUTONG. 2018. *Vödabuk tümologik stamädavödas Volapükik*. Beijing: Vpaklub valemik Tsyinänik.