Positional Asymmetry of Consonants

A Preliminary Study

Cong Luan Tran

BA thesis Linguistics

Under the supervision of

Paul Boersma

Faculty of Humanities
University of Amsterdam
April 23, 2024

Contents

Chapter 1

Introduction

1.1 Motivation

Anderson (2013) claimed that the velar nasal /ŋ/ is exceptional in its phonotactic distribution: virtually all languages that have phonemic /ŋ/ allow it in the final position, while over one-third of those languages prohibit it in the initial position. In fact, of 20 chapters concerning phonology in Dryer & Haspelmath (2013), Anderson's The Velar Nasal is the only chapter that deals with only a single segment. This creates an impression that the distributional asymmetry of /ŋ/ is an exception amongst the consonants of the languages of the world. However, a recent article on the distribution of rhotics also reported a similar asymmetry: in 39% of the languages in the survey, word-initial rhotics are either completely absent or only appear in a low proportion of the lexicon while such restriction is rare in other positions (Labrune 2021). This calls for a question: what about other consonants? Intuitively, some consonants are expected to also have such asymmetry. For example, it should be rare to see /h/in the final position, while some others, like /m/ should be very likely to have a symmetric distribution. Unfortunately, I am unable to find any database that contains such information about the distribution of segments. Thus, this thesis is an attempt to build a small database of consonantal distributions of 20 languages to investigate the degrees of asymmetry of their consonants.

1.2 Literature Review

1.2.1 The Velar Nasal

Anderson (2013) surveyed a total of 469 languages and divided them into three categories: no velar nasal (e.g. French), initial velar nasal (Cantonese), and no initial velar nasal (e.g. English). This categorization seems to imply that languages in the "initial velar nasal" category also allow velar nasal in the final position (because there is no "no final velar nasal" category). However, in further discussion, the author revealed that this is not true. Some languages only have the velar nasal in the initial position but not the final. However, there are two distinct causes for this distribution. The first one is that it is a genuine asymmetry of this one consonant, like in Nenets. The other case is because the language

does not allow any coda at all, like in Fijan. This is an important point to consider since they are two different restrictions. Taking an OT constraints analogy, one is NoCoda constraint and the other is *Coda-ŋ. Should such distinctions appear in the database for this thesis, they ought to be coded differently.

1.2.2 Word-Initial Rhotic Avoidance

Labrune (2021) examined a phenomenon called "Word-initial rhotic avoidance" (WIRA) in a language sample size of 200. The list of languages used is obtained from the 200-language sample recommended in Dryer & Haspelmath (2013). All four terms in the name need clarification to better understand this work.

"Word" is a widely used term but lacks of a definite definition (see discussion in, for example, Haspelmath (2017)). In Labrune (2021), the author adopted a hybrid definition for "word", depending on the distribution of rhotics in the language. The base case is if the rhotics cannot appear in the initial position of any morph, whether prefixes, clitics or roots, then "word" is synonymous to morph. Another case is that a language may allow r-initial prefixes or clitics but not in an independent lexeme. The example given by the author is Japanese, which at first appear to not restricting initial rhotics. However, according to Labrune (2021), all of the initial rhotics are due to "secondary development" like loanword, mimetic word, suffix or initial vowel deletion. Thus, in these cases, "word" is defined as full lexeme, i.e. excluding the secondary development cases. Therefore, the author considered Japanese to be a WIRA language.

"Initial" here means the first segment of a word at the phonemic level, and not necessarily at the phonetic level, and so the author divided WIRA into (phon)emic-WIRA and (phon)etic-WIRA. Emic-WIRA is defined as a language having at least one phonemic rhotic but no rhotic-initial word. Examples for this case include Spanish, Japanese, Basque and Turkish (Labrune 2021). But in some case the avoidance can be considered as happening at the phonetic level, which lead to etic-WIRA. Two situations constitute the etic-WIRA: either a language having initial rhotic at the phonemic level but which turned into non-rhotic at the phonetic level (initial derhoticization), or a phonemic non-rhotic turning into rhotic in other positions except initial (non-initial rhoticization). A wide-known example for non-initial rhoticization is flapping in American English where /t/ and /d/ turned in to [ɾ] in intervocalic position but not in initial. For the initial derhoticization case, one example given by the author is in the language Wichita, where /ɾ/ becomes [n] in the initial position.

"Rhotic" is defined in the paper as liquids minus laterals. The full list is given in Figure 1.1. The uvular rhotics are termed conditional, this means that they will only be considered rhotic if no core rhotics exist in the language and they do not pattern with non-sonorant segments.

"Avoidance" is used instead of "prohibition" because, according to the author, the acceptance of rhotics in a language cannot always be answered with yes or no but is a matter of degree. An example given in the paper is that in a corpus of Maninka with over 28.000 tokens, rhotics appear 1564 times in total, but only 16 times in the initial position. So, the indicator that the author suggested for assessing Avoidance is the ratio between word-initial occurrence of rhotics and the total occurrence, or the ratio between word-initial and non-word-initial

	CORE (UNCONDITIONAL) RHOTICS			SYSTEMIC (CONDITIONAL) RHOTICS
	DENTAL	ALVEOLAR	RETROFLEX	UVULAR
Trills	r̪	r	ɽ	ʀ
Taps or flaps	ɾ̪	ɾ		
Approximants		ɹ	ɻ	ʁ
Lateral flap		ɺ	ɺ̢	

Figure 1.1: Rhotics list (Labrune 2021: p. 4)

rhotics. If the ratio is small, then we can say that the language exhibits WIRA. The result of the paper is that 78 languages exhibit emic-WIRA and 30 languages exhibit etic-WIRA (nine of those also exhibit emic-WIRA). Since I will not investigate corpus in this thesis due to time restriction (and frequency data may not be mentioned in the descriptive grammar), the "avoidance" part is irrelevant to me.

1.2.3 Language Sampling and WALS Sample List

According to Velupillai (2012), there are three types of sampling used in typological studies: probability sampling, variety sampling, and convenience sampling. Probability sampling aims to be statistically representative, and thus tries to achieve genealogical and areal balance in the sample. Variety sampling aims to include as many variations as possible and thus may pay less attention to the genealogical or areal balance. Convenience sampling is, as its name, the sample that the researcher is able to access, though one can still try to be as diverse and representative as possible within it.

The World Atlas of Language Structures (WALS), while allowing the author of each chapter to choose the language sample they wish, provides a 100-language sample and a 200-language sample (built up on the 100-language sample) recommended to be included in every chapter (Comrie et al. 2013). The 100 and 200 languages lists are constructed to maximize the diversity and to be representative of languages of the world, while also concerned with the availability of descriptive data on the languages. The languages of the 100-language sample are more well-known and have more detailed descriptions with better availability (Comrie et al. 2013). Thus, the 100-language sample will be the starting point for constructing the 20-language sample in this thesis using semi-convenience sampling (more will be explained in the methodology).

1.2.4 Property-Driven Phonological Typology

Hyman (2009) claimed that there are two ways of doing typology: one focuses on classifying languages into "types", e.g. tone language and pitch-accent language, and the other focuses on specific properties of the language. The paper is also a criticism of language classification and advocates for focusing on properties instead (i.e. property-driven typology). In this thesis, I will not attempt language classification, that is, I will not put languages into a type like "/ŋ/ symmetric

language" or "more symmetric language" (based on the number of consonants that are symmetric in the language). I will only focus on the properties of each consonant in each language, the interest here are the segments themselves and not the languages that contain them.

Arguably, with this focus on the segment, it could be better to sample the segments and not the languages, i.e. investigate the distribution of 20 segments in all languages instead of all segments in 20 languages. However, the dimensions for segment sampling would be too large. For example, the ideal database for choosing 20 segments to investigate would be PHOIBLE (Moran & McCloy 2019), but for each segment, their presence would be counted in 3020 inventories of 2186 distinct languages. Within the scope of this thesis, that number would be impossible to work with, so I need to choose a more viable approach of sampling based on language.

Chapter 2

Methodology

2.1 Language Sampling

The sampling method I use is semi-convenience. It is "semi-", and not purely convenience, because the languages are chosen randomly from the 100-language sample of WALS. I draw a random 20-language list from the 100-language sample, then skim through the reference grammar of the language given in WALS to see if it contains the information I want. If the reference grammar is not accessible or it does not provide information on the consonantal distribution for the language, I attempt to find the data in other papers about the language. When that also fails, the language is replaced by another language (also chosen randomly from the 100-language sample) from a language family not already presented in the list.

2.2 Data Treatment

2.2.1 Phonetics or Phonology?

In works that consider the representation of phonology and phonetics in great depth like Boersma (1998, 2007), phonology and phonetics is considered to have at least four levels: Underlying and Surface forms of phonology, and Articulatory and Auditory forms of phonetics. Phonetic data, in the sense of acoustic signals and articulatory gestures, are probably the most 'raw' data that can be used, while phonemes are already subjected to the analysis and perspective of the authors who wrote these grammars. Phonemic analysis can produce very different ways of categorizing sounds of a language, which may be justified for the language in question but may not fit for cross-linguistics comparison.

Ideally, I should have gathered data at the phonetic levels to have the primary data, but this task is impossible due to two problems. The first problem is theoretical in that most reference grammars operate on the two levels model of phonetics-phonology: there are the Underlying Form (UF) of phonology, and Surface Form (SF) of phonetics. While their Underlying Form may correspond to Underlying Form in Figure 2.1, the degree of phonetic details in their Surface Form is unclear. The second problem is more practical: most reference grammars do not focus on phonetic details, and so no phonetic data is available to gather
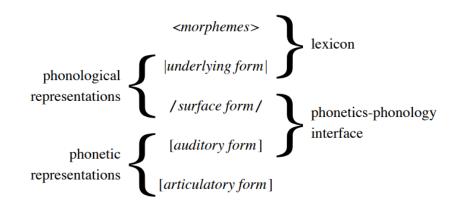
Figure 2.1: Multi-level representation of phonetics and phonology (Boersma 2009)

from them. There are occasionally phonetic remarks, but those are scattered and not made in a systematic manner.

Therefore, I decided to take a practical middle ground and record the data available at the allophonic level. For example, if English happens to be on my list then the 'clear L' [l] and the 'dark L' [ɫ] will be coded as two distinct sounds, with [l] only allowed in the onset, and [ɫ] only in the coda.

2.2.2   Bracket Notation

The data collected and used in this paper resides, to some extent,between the boundary of phonetic and phonemic, so I have to make certain notation decisions. In this paper, whenever I refer to the segments that I processed during data collection, I will use the square bracket '[]', since my aim is to record it at phonetic level. The original reports from reference grammar may vary in terms of phonetic precision, but for the purpose of this paper I will consider them to be equal. However, when referring to the segments that are used in the argument of other authors, I will respect their use of bracket.

2.2.3   Cluster and segmentation

Consonant clusters will not be considered in this thesis and individual segments in a cluster will be counted separately. However, I will respect the segmentation analysis of the author. Thus, if an author insists that, for instance, [ts] is one segment in this language, it will be coded as a sound distinct from both [t] and [s].

2.2.4   Exclusion of certain allophonic variation

Since the purpose of this thesis is to investigate the distributional asymmetry of consonants relative to the syllable, allophonic variations conditioned by other factors (e.g. as-/dis- similation) will be ignored. For example, if in a language, there are [tʃ] and [k] that are in complementary distribution where [tʃ] only appears before front vowels and [k] elsewhere, both of them will be recorded as one segment [k].

This decision is not a statement about the real phonological representation of those segments, but merely a way to avoid confounding factors. An alternative choice is to record both of them separately, but this could disturb the data on positional distribution. For example, [tʃ] may ends up having higher count and especially in the initial position, because it is quite common to see /t/ and /k/ become [tʃ] in the surface form when followed by /i/ or /j/. However, in this case the counting of [tʃ] in the initial position is caused partially by the proximity with a sound that has palatal constriction and not purely its position within a syllable. Arguably, it could also be attributed to the position of the segment (assimilation in the initial position and not in final position), but then we may have to deal with anticipatory versus perseverative assimilation, which is far beyond the scope of this paper.

2.2.5 Data Encoding

Three positions in the syllable will be distinguished: initial, medial, and final. Initial and final means onset and coda, and medial means the intervocalic environment. If a segment is allowed in a position, it will be coded as 1, otherwise 0. However, if a language prohibits all final consonants, the segments will be coded as "N/A" for this position instead.

In my experience, sometimes the author writes nothing about intervocalic phonotactics. So, following the Maximal Onset Principle (MOP), if the author does not provide any information for the medial position, segments in this position will be assumed to share the same distribution with the initial position. This assumption does not mean that I believe MOP to be true, but it is a practical decision to fill in the gap in the data.

Language	Segment	Initial	Medial	Final
A	t	1	1	1
A	d	1	1	0
...				
Z	h	1	0	N/A

Table 2.1: Example table

2.2.6 Data Analysis

For each segment in each position the number of "1" will be divided by the sum of the number of "1" and "0", yielding a restriction percentage where 1.00 means not restricted while 0.00 means completely not allowed.

Chapter 3

Result

3.1 Data

Table 3.1: The List of Languages

code	Name	Genus	Family	Reference
fij	Fijian	Oceanic	Austronesian	Dixon (1988)
tha	Thai	Kam-Tai	Tai-Kadai	Tingsabadh & Abramson (1993)
hau	Hausa	West Chadic	Afro-Asiatic	Newman (2000)
lav	Lavukaleve	Lavukaleve	Solomons East Papuan	Terrill (2011)
qim	Quechua (Imbabura)	Quechuan	Quechuan	Cole (1982)
may	Maybrat	Maybrat	Maybrat	Dol (2007)
imo	Imonda	Border	Border	Seiler (1985)
kio	Kiowa	Kiowa-Tanoan	Kiowa-Tanoan	Watkins & McKenzie (1984)
mnd	Mandarin	Chinese	Sino-Tibetan	Li & Thompson (2009)
spa	Spanish	Romance	Indo-European	Geeslin (2018)
war	Wari'	Chapacura-Wanham	Chapacura-Wanham	Everett & Kern (2002)
kor	Korean	Korean	Korean	Shin, Kiaer & Cha (2012)
vie	Vietnamese	Viet-Muong	Austro-Asiatic	Kirby (2011)
chk	Chukchi	Northern Chukotko-Kamchatkan	Chukotko-Kamchatkan	Dunn (1999)
wra	Warao	Warao	Warao	Romero-Figeroa (2003)
knd	Kannada	Dravidian	Dravidian	Sridhar (1990)
sup	Supyire	Senufo	Niger-Congo	Carlson (1994)
tiw	Tiwi	Tiwian	Tiwian	Osborne (1974)
lan	Lango	Western Nilotic	Eastern Sudanic	Noonan (1992)
kew	Kewa	Enga_Kewa-Huli	Trans-New Guinea	Franklin (1971)

The 20 languages used in this survey are presented in Table 3.1 along with their genealogical data and the reference used. A total of 411 tokens of 95 different segments was recorded, the 20 most common segments and their restriction percentages are presented in Table 3.2. The full list of all segments is provided in Appendix B.

3.2 Overview

The list of 20 most frequent segments in this study is similar to the top 20 segments list reported in Gordon (2016: p. 45), which comes from a survey of 317 languages. The only two different segments are [ʃ] and [ɲ] in his list compared

9

Table 3.2: List of most common segments

Segment	Count	Initial	Medial	Final
m	20	1.00	1.00	0.94
n	20	1.00	1.00	0.94
t	20	1.00	0.90	0.56
k	19	1.00	0.95	0.47
p	18	0.94	0.89	0.43
s	18	0.94	1.00	0.57
l	16	0.94	0.94	0.62
j	15	0.93	0.93	0.55
d	13	0.92	0.85	0.20
w	13	1.00	1.00	0.56
b	12	0.92	0.92	0.30
ŋ	12	0.67	0.67	1.00
g	11	0.82	0.91	0.22
f	10	1.00	0.90	0.63
h	10	1.00	0.90	0.11
r	10	0.90	1.00	0.57
ɾ	8	0.63	1.00	0.71
tʰ	8	0.88	0.88	0.25
tʃ	7	1.00	1.00	0.00
ʔ	7	0.86	1.00	0.40

to [r] and [tʰ] in my list. There are even some resemblance in terms of the frequency of each segments: [n t m k] are the four most common segments in both lists. Therefore, despite the small scale of this study the list of top 20 segments can still be considered representative of world languages to some degree.

For most segments in this survey, we can see that the distributional discrepancies between different positions are apparent. The most symmetric segments are [n] and [m], which have equal occurrences in initial and medial positions (20 languages), with slightly fewer instances in the final position (19 languages each). The distribution of velar nasal [ŋ] is consistent with the finding of Anderson (2013), namely that all languages that have [ŋ] allow it in the final position while only two-thirds of them allow it in the initial and medial position. Two most common "rhotics" in the data, [r] and [ɾ] also exhibit the Word-Initial Rhotic Avoidance to some degree, although it is more apparent for [ɾ] (63% compared to 90% of [r]).

The non-continuant segments (i.e. with the oral tract completely blocked) in this study can be divided into three main groups: nasal, voiced stop and voiceless stop. It may be useful to discuss these segments together because all three major places of articulation (labial, coronal, and dorsal) of them appear in the top 20 most common segments of the study. In general, the nasal segments seem to be more symmetric than the voiceless, which in turn are more symmetric than the voiced. The discrepancies lie mainly on the final position, since acceptance of them (except [ŋ]) is nearly 100% in initial and medial position. The voiced stop is the most restricted group in this position, only allowed in under one third of the

languages in which they occur. They are followed by the voiceless stops, which is allowed in the final position for around a half of the languages that they appear in. The least restricted are the nasals, which are near universal also in the final position (100% for [ŋ], and 94% for each of [m] and [n]).

Turning to the fricatives, we find three fricative segments in Table 3.2: [s], [f] and [h]. There is a large difference in the frequency of [s], which occurs in 18 languages compared to the other two fricatives, which occur in only 10 languages each. The behaviour of [s] and [f] are fairly similar to the voiceless stops. They are also near universal in the initial and medial position and around 50% in the final position However, [s] appears slightly more favoured in the medial position than initial position. The glottal fricative [h], on the other hand, is much more restricted in the final position (only 11%) while still having nearly universal acceptance in the other two positions.

The non-nasal sonorants in Table 3.2 are comprised of [l j w r ɾ]. Most of them behave similar to the voiceless stops, with near universal presence in the initial and medial, and around 50% acceptance in final position. The exception is the rhotic tap [ɾ] which is much more restricted in the initial position while still universally accepted in the medial position. In fact initial is the most restricted position of [ɾ], even more than the final position. The final position of [ɾ], on the other hand, has the highest acceptance out of all non-nasal segments (71%).

The remaining miscellaneous segments are the aspirated stop [tʰ], the glottal stop [ʔ], and the affricate [tʃ]. The affricate is perhaps the most special segment out of the three. It is the only segment in Table 3.2 that is totally inhibited from occurring in the final position. The glottal stop behaves similar to other (plain) voiceless stops while the voiceless aspirated [tʰ] behaves more like voiced stops. Although also being voiceless, [tʰ] has lower acceptance than all other voiceless stops in all position, including the final.

3.3 Discussion

The non-continuants [m n ŋ p t k b d g] are very frequent in languages of the world, and as seen in Table 3.2, all of these segments also appear amongst the most common list. So, I believe that we can start or discussion with these segments and their distribution. I first discuss about the discrepancies between the onset-coda acceptance of nasals and obstruents and speculate some possible explanation. Then, I will look at the distribution of the voiceless and voiced stop.

The asymmetry between the set of segments that can occur in the onset and coda are well-known to linguists. In the tradition of Optimality Theory (OT), there are two major markedness constraints regards syllable structure: ONSET and NOCODA. The first of them, ONSET, requires a syllable to have a consonantal onset, which is violated by vowel-initial syllable. The NOCODA constraint states that a syllable should not end with a consonant. Thus, the syllables that have coda will violate the constraint and become less optimal than syllables without it. According to OT, only the optimal form, meaning form that violates the least constraints, will be chosen to produce by the speaker. Only when the faithfulness constraint is ranked above the markedness constraint that a syllable with coda in the underlying form can have its coda produced. Therefore, it can be said that

the 'default' option for syllable structure in languages is without a coda. Gordon (2016) also reported a similar observation, that if a language permits syllables with coda, it will also permit syllables without coda.

But it is clear from the data that not all coda are treated equally. If a language allows its syllable to have coda, the three nasals are almost universally a candidate for that position, while other segments are generally much less favoured. Not all segments violate the NOCODA constraint to the same degree, and there is some attribute of the nasals that make them violate the constraint to a lesser extend. In Clements (1990), it is proposed that the syllable structure follows a Sonority Sequencing Principle, which states that the syllable typically only allows one sonority peak, usually a vowel, with sequences of consonants on their sides. This principle also requires that the segments in the sequences have to be ordered, in term of sonority, in decreasing order toward the edges. The author observed that the initial and final of a syllable follow different generalization: while segments in the onset prefers to maximize their sonority difference from the nucleus, segments in the coda prefer minimizing their sonority difference. This means that if a language permits any coda at all, then the more sonorant codas (like nasals) are more preferred than their less sonorant counterparts (like stops).

The difference in the restriction of voiced and voiceless stops can be expected by the well-known Final Obstruent Devoicing process. Final Obstruent Devoicing means that voiced stops can be neutralized to their voiceless counterpart when they are at the end of the word. Another related phenomena is Intervocalic Voicing where voiceless stop becomes voiced when it is between vowels. However, the predicted effect of this phenomena, namely that voiceless stops are more restricted than voiced stops in medial position, is not observed in the top 20 segments. The likelihood of voiced and voiceless segments in medial position is found to be fairly equal. This is quite surprising because some accounts consider them to be similar phonological processes, namely weakening processes (Gordon 2016, Harris 2009). Both can be seen as reducing the articulation effort: both abducting the vocal folds between two vowels to maintain voicing at the end of the word takes more effort (Gordon 2016). With these similarities between the two processes, it seems surprising why the effect of one is observed and the other is not. Thus, the account of a weakening process is not enough to explain the distribution of voiced and voiceless segments in this data.

It is possible that the efforts in the two cases are not the same and one has more priority than the other. Indeed, voicing in vowels and consonants, although usually represented in phonology by the same feature, are different in terms of aerodynamic. The voicing of vowels and sonorants is due to an absence of built-up air pressure in the oral tract while the voicing of obstruent requires such air pressure (Harris 2009). Thus, even when intervocalic voicing avoiding changing the adduction of the vocal folds (which is required for articulating vowels), it still has to increase the air pressure to make obstruent voicing possible. So, intervocalic voicing can be seen as choosing between two efforts, changing the states of the vocal folds or increasing air pressure, while final devoicing only minimize efforts. That might be the reason why the preference for voicelessness of the stop in the final position is better reflected in the data.

3.4 Interesting Findings

3.4.1 The Glottal Stop

The most phonetically favourable position for glottal stop is probably word-initial. Even languages like English, where glottal stop does not have phonemic status, still seem to favour word-initial glottal stop at the phonetic level (Garellek 2012). In fact, Garellek (2012) went as far as to suggest that perhaps all languages that does not contrast /#V/ and /#ʔV/ could insert [ʔ] at the beginning of the word, though they may do so with different degrees of frequency. Consider this in the context of the basics OT constraint ONSET, the surface forms with [ʔ] in the onset would outrank the vowel-initial forms. Therefore, if a glottal stop exists in a language, it should be allowed in the initial position.

However, the data gathered in Table 3.2 shows that only seven of of the eight languages with glottal stop allow it in the initial position. The only language that does not allow initial glottal stop is Supyire, a language of the Niger-Congo family. This language's syllable structure is CV or CVV, with the exception of a few grammatical words that allow word-initial vowel (Carlson 1994: p. 7). Given such syllable structure, one may expect that if the glottal stop is recognized in the language, its most favourable position would be word-initial. However, the glottal stop is reported to be only a marginal phoneme in Supyire and can only appear in intervocalic environment. Carlson (1994) speculated, based on evidence from loan words, that /ʔ/ (the author uses the symbol /h/) might be a reduction of an earlier /g/. But because in the current phonological system of the language there is also a [g] that can appear in intervocalic position, I consider [ʔ] to be considered a separate phoneme instead of an allophone of [g].

3.4.2 Final Avoidance of Two Nasals

Although the nasals are supposed to be well-accepted in the final position, and the velar nasal is universally accepted, Table 3.2 still shows two cases where /n/ or /m/ is not allowed in the coda. The first of them, the absence of /m/ in the coda, is observed in Mandarin Chinese. The second one is the absence of final /n/, which is observed in Hausa, an Afro-Asiatic language. The reason for the absence is different in each case: for Mandarin it is because of a historical merging process, while for Hausa it is allophony.

The rhyme table for Mandarin Chinese presented in Li & Thompson (2009) contains only syllables that end with either a vowel or one of the two nasals [n] and [ŋ]. The labial nasal [m], despite being a legit segment in the onset, does not appear in the coda. More surprisingly, /m/ is a legit segment in the coda in Middle Chinese, and still remain so in a few other Sinitic languages like Cantonese (Zee 1985). The explanation given by Zee (1985) is that the labial nasal /m/ has merged into /n/ in Mandarin Chinese. Middle Chinese allowed all three nasals /m/, /n/ and /ŋ/ to appear in the end of the word, but there is a merging process that affects all Sinitic languages. In this process, the labial nasal /m/ seems to be the prime target for being merged, and in Mandarin it is merged with /n/ (Zee 1985).

For the case of Hausa, it is more due to methodological decision made in

the process of data gathering. Specifically, Newman (2000) presented Hausa as having only two phonemic nasals: /n/ and /m/. But he also noted that /n/ in the final position is always realized as [ŋ] and never as [n]. Therefore, in my data gathering Hausa is considered having both [n] and [ŋ], with [n] allowed in initial and medial and [ŋ] allowed in only final position. The labial /m/ is also usually realized as [ŋ], but there are some cases where it is still realized as [m]. Thus, [m] is still coded as being allowed in all three position in my data.

Thus, the absence of word-final /m/ in Mandarin and /n/ in Hausa is not an evidence against the universal acceptance of nasals in word-final position. They are some particular cases of historical changes or allophony within the space of the nasal inventory. But another interesting question that they pose is "Why the velar nasal?". Although in Mandarin /n/ still remain distinguished from /ŋ/, some other Sinitic languages also show a process of /n/ merging to /ŋ/ Zee (1985). Combined that with the distribution of /ŋ/ in word-final position, as observed in Anderson (2013), it seems like the velar nasal holds a strong preference in this position. One possible speculation is that the velar nasal is the most vowel-like nasal. And if it is true that languages prefer to end words with a vowel, then the velar nasal could be the most highly ranked alternative.

Chapter 4

Conclusion

Despite the small language list used in this study, and thus the lack of significant statistical analysis, a few patterns can be observed from the data of the 20 most common segments. We can first conclude that asymmetry is indeed the norm for most segments, with the exception of two nasals /n/ and /m/. The nasals, if they exist in a language, are always allowed in the final position. Two of them, /m/ and /n/, are also always allowed in the initial and medial position. There are two cases where /m/ and /n/ are absent from the final position but they are due to a process of merging or allophony with another common nasals. The velar nasal /ŋ/, on the other hand, is only universally allowed in the final position and is more restricted in other position. While this observation conforms to the report from Anderson 2013, neither the fact that the distribution of /ŋ/ is asymmetric nor its degree of asymmetry is special. However, what makes the distribution of /ŋ/ special is that it is the only segment that is skewed toward the final position.

One particular question that remains is why the nasals are that well-accepted in the final position, and a possible reason could be due to its high sonority. But that answer is still unsatisfying since other sonorants like liquids and glides are just as popular in the final position as the voiceless stops, which is on the other end of the sonority scale. Perhaps the nasals are more similar to vowels than other sonorants in some attributes that are not considered in this thesis.

Turning to the voiced and voiceless stops, the strong restriction on the voiced stops in the final position is well-expected by the Final Obstruent Avoidance process. However, the effect of the related phenomena of intervocalic voicing is not observed in the data. A speculated explanation is that maintaining voicing in word-final position takes more effort than keeping the stops voiceless between two vowels. Therefore, the need to avoid voicing in word-final stops is more profound and reflected better in the data.

Due to numerous limitation, many explanations of the data in this study have to remain at the speculation stage. Nevertheless, it can serve as a preliminary research to look for direction for further studies. For example, another research could try to compare the similarity between vowels and nasals versus vowels and other sonorants as a way to explain the high acceptance of nasals in word-final position.

Appendix A

Full Distribution Table

Language	Segment	Initial	Medial	Final
fij	m	1	1	N/A
fij	p	1	1	N/A
fij	mb	1	1	N/A
fij	β	1	1	N/A
fij	f	1	1	N/A
fij	n	1	1	N/A
fij	t	1	1	N/A
fij	nd	1	1	N/A
fij	ð	1	1	N/A
fij	r	1	1	N/A
fij	nr	1	1	N/A
fij	s	1	1	N/A
fij	tʃ	1	1	N/A
fij	l	1	1	N/A
fij	j	1	1	N/A
fij	ŋ	1	1	N/A
fij	k	1	1	N/A
fij	ŋg	1	1	N/A
fij	w	1	1	N/A
fij	ʔ	1	1	N/A
lav	pʰ	1	1	0
lav	tʰ	1	1	1
lav	t	1	0	0
lav	kʰ	1	1	1
lav	b	1	1	0
lav	mb	0	1	0
lav	d	1	1	0

lav	nd	0	1	0
lav	m	1	1	1
lav	n	1	1	1
lav	ŋ	1	1	1
lav	r	1	1	1
lav	l	1	1	1
lav	f	1	1	1
lav	s	1	1	1
lav	h	1	1	0
lav	β	1	1	1
lav	ɰ	1	1	1
qim	p	1	1	0
qim	t	1	1	0
qim	k	1	1	0
qim	ts	1	1	0
qim	tʃ	1	1	0
qim	b	1	1	0
qim	d	1	1	0
qim	g	1	1	0
qim	ɸ	1	1	0
qim	s	1	1	1
qim	ʃ	1	1	1
qim	x	1	1	1
qim	β	1	1	0
qim	z	1	1	0
qim	ʒ	1	1	0
qim	m	1	1	1
qim	n	1	1	1
qim	ɲ	1	1	0
qim	ŋ	0	0	1
qim	l	1	1	1
qim	ʐ	1	0	0
qim	ɾ	0	1	1
qim	w	1	1	1
qim	j	1	1	1
may	p	1	1	0
may	t	1	1	1
may	tʰ	0	0	1

17

may	tˀ	0	0	1
may	k	1	1	1
may	g	0	1	0
may	kˀ	0	0	1
may	m	1	1	1
may	n	1	1	1
may	f	1	1	1
may	s	1	1	1
may	x	1	1	1
may	r	1	1	1
may	ɾ	0	1	1
may	w	1	1	0
may	j	1	1	0
imo	f	1	0	1
imo	v	0	1	0
imo	p	1	1	0
imo	b	1	0	0
imo	mb	0	1	1
imo	m	1	1	1
imo	n	1	1	1
imo	s	1	1	1
imo	l	1	1	1
imo	t	1	1	1
imo	d	1	0	0
imo	nd	0	1	1
imo	k	1	1	1
imo	g	1	0	0
imo	ŋg	0	1	1
imo	h	1	1	1
imo	ɦ	0	1	0
kio	p	1	1	1
kio	p'	1	1	0
kio	pʰ	1	1	0
kio	b	1	1	0
kio	m	1	1	1
kio	t	1	1	1
kio	t'	1	1	0
kio	tʰ	1	1	0

kio	d	1	1	0
kio	n	1	1	1
kio	l	1	1	0
kio	ᵈl	0	0	1
kio	ts	1	1	0
kio	ts'	1	1	0
kio	s	1	1	0
kio	z	1	1	0
kio	j	1	1	1
kio	k	1	1	0
kio	k'	1	1	0
kio	kʰ	1	1	0
kio	g	1	1	0
kio	h	1	1	0
mnd	p	1	1	0
mnd	pʰ	1	1	0
mnd	t	1	1	0
mnd	tʰ	1	1	0
mnd	k	1	1	0
mnd	kʰ	1	1	0
mnd	ts	1	1	0
mnd	tʂ	1	1	0
mnd	tɕ	1	1	0
mnd	tsʰ	1	1	0
mnd	tʂʰ	1	1	0
mnd	tɕʰ	1	1	0
mnd	m	1	1	0
mnd	n	1	1	1
mnd	f	1	1	0
mnd	s	1	1	0
mnd	ʂ	1	1	0
mnd	ɕ	1	1	0
mnd	x	1	1	0
mnd	l	1	1	0
mnd	ɹ	1	1	0
mnd	ŋ	0	0	1
spa	p	1	1	1
spa	b	1	1	1

spa	m	1	1	1
spa	f	1	1	1
spa	θ	1	1	1
spa	t	1	1	1
spa	d	1	1	1
spa	s	1	1	1
spa	n	1	1	1
spa	l	1	1	1
spa	ɾ	1	1	1
spa	r	1	1	1
spa	tʃ	1	1	0
spa	ɲ	1	1	0
spa	ɟ	1	1	1
spa	ʎ	1	1	0
spa	k	1	1	1
spa	g	1	1	1
spa	x	1	1	1
war	p	1	1	0
war	pˀ	0	0	1
war	t	1	1	0
war	tˀ	0	0	1
war	tʙ̥	1	1	0
war	k	1	1	0
war	kˀ	0	0	1
war	kʷ	1	1	0
war	ʔ	1	1	1
war	tʃ	1	1	0
war	h	1	1	0
war	m	1	1	1
war	mb	1	1	0
war	mʔ	0	0	1
war	n	1	1	1
war	nd	1	1	0
war	nʔ	0	0	1
war	ɾ	1	1	0
war	w	1	1	0
war	j	1	1	0
kor	p	1	0	0

kor	p̬	1	1	0
kor	pʰ	1	1	0
kor	p˼	0	0	1
kor	b	0	1	0
kor	t	1	0	0
kor	t̬	1	1	0
kor	tʰ	1	1	0
kor	t˼	0	0	1
kor	d	0	1	0
kor	k	1	0	0
kor	k̬	1	1	0
kor	kʰ	1	1	0
kor	k˼	0	0	1
kor	ɡ	0	1	0
kor	s	1	1	0
kor	s̬	1	1	0
kor	h	1	0	0
kor	ɦ	0	1	0
kor	tɕ	1	0	0
kor	dʑ	0	1	0
kor	tɕ̬	1	1	0
kor	tɕʰ	1	1	0
kor	m	1	1	1
kor	n	1	1	1
kor	ŋ	0	0	1
kor	l	0	0	1
kor	ɾ	1	1	0
vie	ɓ	1	1	0
vie	p	0	0	1
vie	m	1	1	1
vie	w	1	1	1
vie	f	1	1	0
vie	v	1	1	0
vie	t	1	1	1
vie	tʰ	1	1	0
vie	n	1	1	1
vie	l	1	1	0
vie	ɗ	1	1	0

vie	s	1	1	0
vie	z	1	1	0
vie	tɕ	1	1	0
vie	ɲ	1	1	0
vie	j	0	0	1
vie	k	1	1	1
vie	ŋ	1	1	1
vie	x	1	1	0
vie	ɣ	1	1	0
vie	ʔ	1	1	0
vie	h	1	1	0
chk	p	1	1	1
chk	m	1	1	1
chk	w	1	1	1
chk	t	1	1	1
chk	n	1	1	1
chk	ɾ	1	1	1
chk	s	1	1	1
chk	ɬ	1	1	1
chk	j	1	1	1
chk	k	1	1	1
chk	ŋ	1	1	1
chk	ɣ	1	1	1
chk	q	1	1	1
wra	p	1	1	N/A
wra	m	1	1	N/A
wra	w	1	1	N/A
wra	t	1	1	N/A
wra	s	1	1	N/A
wra	r	0	1	N/A
wra	d	1	0	N/A
wra	n	1	1	N/A
wra	j	1	1	N/A
wra	k	1	1	N/A
wra	kʷ	1	1	N/A
wra	h	1	1	N/A
knd	p	1	1	0
knd	pʰ	1	1	0

knd	b	1	1	0
knd	bʰ	1	1	0
knd	t	1	1	0
knd	tʰ	1	1	0
knd	d	1	1	0
knd	dʰ	1	1	0
knd	ʈ	1	1	0
knd	ʈʰ	1	1	0
knd	ɖ	1	1	0
knd	ɖʰ	0	1	0
knd	k	1	1	0
knd	kʰ	1	1	0
knd	g	1	1	0
knd	gʰ	1	1	0
knd	tʃ	1	1	0
knd	dʒ	1	1	1
knd	f	1	1	1
knd	s	1	1	1
knd	z	1	1	0
knd	ʃ	1	1	0
knd	ʂ	1	1	0
knd	m	1	1	1
knd	n	1	1	1
knd	ɳ	0	1	0
knd	l	1	1	0
knd	ɭ	1	1	0
knd	r	1	1	0
knd	j	1	1	0
knd	v	1	1	0
sup	p	1	1	N/A
sup	b	1	1	N/A
sup	f	1	1	N/A
sup	v	1	1	N/A
sup	m	1	1	N/A
sup	t	1	1	N/A
sup	d	1	1	N/A
sup	ɾ	0	1	N/A
sup	s	1	1	N/A

sup	z	1	1	N/A
sup	n	1	1	N/A
sup	l	1	1	N/A
sup	tʃ	1	1	N/A
sup	dʒ	1	1	N/A
sup	ʃ	1	1	N/A
sup	ʒ	1	1	N/A
sup	ɲ	1	1	N/A
sup	j	1	1	N/A
sup	k	1	1	N/A
sup	g	1	1	N/A
sup	ʀ	0	1	N/A
sup	ŋ	1	1	N/A
sup	w	1	1	N/A
sup	ʔ	0	1	N/A
tiw	p	1	1	0
tiw	m	1	1	1
tiw	t̪	1	1	0
tiw	n̪	1	1	1
tiw	j	1	1	0
tiw	t	1	1	0
tiw	n	1	1	1
tiw	l	1	1	1
tiw	r	1	1	0
tiw	ɹ	1	1	1
tiw	k	1	1	0
tiw	ŋ	1	1	1
tiw	ɣ	1	1	0
tiw	w	1	1	0
lan	p	1	1	1
lan	t	1	1	1
lan	tɕ	1	1	1
lan	k	1	1	1
lan	b	1	1	1
lan	d	1	1	1
lan	dʑ	1	1	1
lan	g	1	1	1
lan	m	1	1	1

lan	n	1	1	1
lan	ɲ	1	1	1
lan	ŋ	1	1	1
lan	ɸ	0	1	0
lan	s	0	1	0
lan	ɕ	0	1	0
lan	x	0	1	0
lan	ɺ̥	0	1	0
lan	ɾ	1	1	1
lan	l	1	1	1
lan	j	1	1	0
lan	w	1	1	0
lan	ʔ	1	1	0
kew	p	1	1	N/A
kew	t	1	1	N/A
kew	k	1	1	N/A
kew	b	1	1	N/A
kew	d	1	1	N/A
kew	g	1	1	N/A
kew	m	1	1	N/A
kew	n	1	1	N/A
kew	l	1	1	N/A
kew	r	1	1	N/A
kew	s	1	1	N/A
kew	w	1	1	N/A
kew	j	1	1	N/A
tha	p	1	1	1
tha	b	1	1	0
tha	pʰ	1	1	0
tha	t	1	1	1
tha	d	1	1	0
tha	tʰ	1	1	0
tha	k	1	1	1
tha	kʰ	1	1	0
tha	ʔ	1	1	1
tha	m	1	1	1
tha	n	1	1	1
tha	ŋ	1	1	1

tha	f	1	1	0
tha	s	1	1	0
tha	h	1	1	0
tha	tɕ	1	1	0
tha	tɕʰ	1	1	0
tha	r	1	1	0
tha	j	1	1	1
tha	w	1	1	1
tha	l	1	1	0
tha	h	1	1	0
hau	b	1	1	1
hau	ɓ	1	1	1
hau	m	1	1	1
hau	ɸ	1	1	1
hau	w	1	1	1
hau	t	1	1	1
hau	d	1	1	0
hau	ts'	1	1	0
hau	ɗ	1	1	0
hau	n	1	1	0
hau	ŋ	0	0	1
hau	s	1	1	1
hau	z	1	1	1
hau	r	1	1	1
hau	ɽ	1	1	1
hau	l	1	1	1
hau	tʃ	1	1	0
hau	dʒ	1	1	0
hau	ʔʲ	1	1	0
hau	ʃ	1	1	0
hau	j	1	1	1
hau	k	1	1	0
hau	g	1	1	0
hau	k'	1	1	0
hau	kʲ	1	1	0
hau	gʲ	1	1	0
hau	kʲ'	1	1	0
hau	kʷ	1	1	0

hau	gʷ	1	1	0
hau	kʷ'	1	1	0
hau	ʔ	1	1	0
hau	h	1	1	0

Appendix B

Full List of Segments

Segment	Count	Initial	Medial	Final
m	20	1.00	1.00	0.94
n	20	1.00	1.00	0.94
t	20	1.00	0.90	0.56
k	19	1.00	0.95	0.47
p	18	0.94	0.89	0.43
s	18	0.94	1.00	0.57
l	16	0.94	0.94	0.62
j	15	0.93	0.93	0.55
d	13	0.92	0.85	0.20
w	13	1.00	1.00	0.56
b	12	0.92	0.92	0.30
ŋ	12	0.67	0.67	1.00
g	11	0.82	0.91	0.22
f	10	1.00	0.90	0.63
h	10	1.00	0.90	0.11
r	10	0.90	1.00	0.57
ɾ	8	0.63	1.00	0.71
tʰ	8	0.88	0.88	0.25
tʃ	7	1.00	1.00	0.00
ʔ	7	0.86	1.00	0.40
kʰ	6	1.00	1.00	0.17
pʰ	6	1.00	1.00	0.00
x	6	0.83	1.00	0.50
z	6	1.00	1.00	0.20
ɲ	5	1.00	1.00	0.25
tɕ	5	1.00	0.80	0.20
mb	4	0.50	1.00	0.33

ʃ	4	1.00	1.00	0.33
v	4	0.75	1.00	0.00
dʒ	3	1.00	1.00	0.50
ɣ	3	1.00	1.00	0.33
k˺	3	0.00	0.00	1.00
kʷ	3	1.00	1.00	0.00
nd	3	0.33	1.00	0.50
ɸ	3	0.67	1.00	0.33
t˺	3	0.00	0.00	1.00
tɕʰ	3	1.00	1.00	0.00
ts	3	1.00	1.00	0.00
β	3	1.00	1.00	0.50
ɓ	2	1.00	1.00	0.50
ɕ	2	0.50	1.00	0.00
ɗ	2	1.00	1.00	0.00
dʑ	2	0.50	1.00	0.50
ɦ	2	0.00	1.00	0.00
k'	2	1.00	1.00	0.00
ŋg	2	0.50	1.00	1.00
p˺	2	0.00	0.00	1.00
ɹ	2	1.00	1.00	0.50
ʂ	2	1.00	1.00	0.00
ts'	2	1.00	1.00	0.00
ʒ	2	1.00	1.00	0.00
bʰ	1	1.00	1.00	0.00
ð	1	1.00	1.00	N/A
dʰ	1	1.00	1.00	0.00
ᵈl	1	0.00	0.00	1.00
ɖ	1	1.00	1.00	0.00
ɖʰ	1	0.00	1.00	0.00
gʰ	1	1.00	1.00	0.00
gʲ	1	1.00	1.00	0.00
gʷ	1	1.00	1.00	0.00
ɟ	1	1.00	1.00	1.00
k̬	1	1.00	1.00	0.00
kʲ	1	1.00	1.00	0.00
kʲ'	1	1.00	1.00	0.00
kʷ'	1	1.00	1.00	0.00

ɭ	1	1.00	1.00	0.00
ɫ	1	1.00	1.00	1.00
ɰ	1	1.00	1.00	1.00
mʔ	1	0.00	0.00	1.00
ɳ	1	0.00	1.00	0.00
n̥	1	1.00	1.00	1.00
nd	1	1.00	1.00	0.00
ⁿr	1	1.00	1.00	N/A
nʔ	1	0.00	0.00	1.00
p'	1	1.00	1.00	0.00
p̥	1	1.00	1.00	0.00
q	1	1.00	1.00	1.00
ʀ	1	0.00	1.00	N/A
ɽ	1	1.00	1.00	1.00
ɾ̥	1	0.00	1.00	0.00
ʂ̥	1	1.00	1.00	0.00
t'	1	1.00	1.00	0.00
ʈ	1	1.00	1.00	0.00
t̪	1	1.00	1.00	0.00
t̥	1	1.00	1.00	0.00
tʙ̥	1	1.00	1.00	0.00
tɕ̥	1	1.00	1.00	0.00
tʰ	1	1.00	1.00	0.00
tʂ	1	1.00	1.00	0.00
tsʰ	1	1.00	1.00	0.00
tʂʰ	1	1.00	1.00	0.00
ʎ	1	1.00	1.00	0.00
ʐ̩	1	1.00	0.00	0.00
ʔʲ	1	1.00	1.00	0.00
θ	1	1.00	1.00	1.00
Total	95			

Bibliography

Anderson, Gregory D.S. 2013. The velar nasal. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. `https://wals.info/chapter/9`.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives* (LOT 11). Zusammenfassung auf niederländ. 501 Gezählt als: NUGI ; 941. The Hague: Holland Academic Graphics. 493 pp.

Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in french. *Lingua* 117(12). 1989–2054. `https://doi.org/10.1016/j.lingua.2006.11.004`.

Boersma, Paul. 2009. Cue constraints and their interactions in phonological perception and production. In *Phonology in Perception*, 55–110. DE GRUYTER MOUTON. `https://doi.org/10.1515/9783110219234.55`.

Carlson, Robert. 1994. *A Grammar of Supyire.* Berlin, New York: De Gruyter Mouton. `https://doi.org/10.1515/9783110883053`.

Clements, George Nickerson. 1990. The role of the sonority cycle in core syllabification. In Beckman ME Kingston J (ed.), *Papers in Laboratory Phonology*, 283–333. Cambridge University Press. `https://doi.org/10.1017/cbo9780511627736.017`.

Cole, Peter. 1982. *Imbabura Quechua* (Lingua Descriptive Studies 5). Asmterdam: North-Holland.

Comrie, Bernard, Matthew S. Dryer, David Gil & Martin Haspelmath. 2013. Introduction. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. `https://wals.info/chapter/s1`.

Dixon, Robert Malcolm Ward. 1988. *A Grammar of Boumaa Fijian.* Chicago, London: University of Chicago Press.

Dol, Philomena Hedwig. 2007. *A grammar of Maybrat: a language of the Bird's Head Peninsula, Papua Province, Indonesia* (Pacific Linguistics 586). Canberra: Australian National University.

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *Wals online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. `https://wals.info/`.

Dunn, Michael John. 1999. *A grammar of Chukchi.* Australian National University dissertation. `https://openresearch-repository.anu.edu.au/handle/1885/10769` (14 October, 2022).

Everett, Daniel L. & Barbara Kern. 2002. *Wari.* London, New York: Routledge. `https://doi.org/10.4324/9780203193327`.

Franklin, Karl J. 1971. *A grammar of Kewa, New Guinea* (Pacific Linguistics 16). Canberra: The Australian National University.

Garellek, Marc. 2012. WPP, No.110: Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics. *UCLA: Department of Linguistics.*

Geeslin, Kimberly L. (ed.). 2018. *The Cambridge Handbook of Spanish Linguistics.* 1st edn. Cambridge: Cambridge University Press. `https://doi.org/10.1017/9781316779194`.

Gordon, Matthew Kelly. 2016. *Phonological typology.* Oxford University Press.

Harris, John. 2009. Why final obstruent devoicing is weakening. In Kuniya Nasukawa & Phillip Backley (eds.), *Strength relations in phonology.* de Gruyter.

Haspelmath, Martin. 2017. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia linguistica* 51(s1000). 31–80.

Hyman, Larry M. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language sciences* 31(2-3). 213–238.

Kirby, James P. 2011. Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41(3). 381–392. `https://doi.org/10.1017/S0025100311000181`.

Labrune, Laurence. 2021. Word-initial rhotic avoidance: a typological survey. *Glossa: a journal of general linguistics* 6(1).

Li, Charles N. & Sandra A. Thompson. 2009. *Mandarin Chinese: a functional reference grammar.* 1st edn. Berkeley: University of California Press.

Moran, Steven & Daniel McCloy (eds.). 2019. *Phoible 2.0.* Jena: Max Planck Institute for the Science of Human History. `https://phoible.org/`.

Newman, Paul. 2000. *The Hausa language: an encyclopedic reference grammar* (Yale Language Series). New Haven: Yale University Press.

Noonan, Michael. 1992. *A grammar of Lango* (Mouton Grammar Library 7). Berlin, New York: Mouton de Gruyter. 352 pp.

Osborne, C. R. 1974. *The Tiwi language: grammar, myths and dictionary of the Tiwi language spoken on Melville and Bathurst Islands, northern Australia* (Australian Aboriginal Studies 55). Canberra: Australian Institute of Aboriginal Studies.

Romero-Figeroa, Andrés. 2003. *A reference grammar of Warao.* 2nd edn. (LINCOM Studies in Native American Linguistics 6). München: LINCOM Europa.

Seiler, Walter. 1985. *Imonda, a Papuan language* (Pacific Linguistics 93). Canberra: Australian National University.

Shin, Jiyoung, Jieun Kiaer & Jaeeun Cha. 2012. *The Sounds of Korean.* 1st edn. Cambridge: Cambridge University Press. `https://doi.org/10.1017/CBO9781139342858`.

Sridhar, S. N. 1990. *Kannada* (Descriptive Grammars). London, New York: Routledge.

Terrill, A. 2011. *A grammar of lavukaleve* (Mouton Grammar Library). Berlin, New York: De Gruyter Mouton.

Tingsabadh, M. R. Kalaya & Arthur S. Abramson. 1993. Thai. *Journal of the International Phonetic Association* 23(1). 24–28. `https://doi.org/10.1017/S0025100300004746`.

Velupillai, Viveka. 2012. *An introduction to linguistic typology.* John Benjamins Publishing.

Watkins, Laurel J. & Parker McKenzie. 1984. *A grammar of Kiowa* (Studies in the anthropology of North American Indians). Lincoln: University of Nebraska Press.

Zee, Eric. 1985. Sound change in syllable final nasal consonants in chinese. *Journal of Chinese linguistics* 13(2). 291–330.