Consonants and Syllable Structure in Angaité (Enlhet-Enenlhet)

APPROVED BY

SUPERVISING COMMITTEE:

Patience Epps, Supervisor

Anthony C. Woodbury

Consonants and Syllable Structure in Angaité (Enlhet-Enenlhet)

by

Paige Erin Wheeler

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

May, 2020

Acknowledgments

The list of people I could thank here is endless, and in the midst of the COVID-19 outbreak, I'm more grateful than ever for my excellent support system. First and foremost, I'm grateful to my consultants and hosts in La Patria this summer, without whom this research would not have been possible—for struggling through not one but two new languages with me this summer, aguije. Thanks to Ben, José, Olivia, Marcelo, and Paula for bringing me much needed study breaks and being familiar faces while I was far from home. To everyone in 'Field 2019!' for their concern and advice while in the field, without which I would have been unproductive and adrift. Thanks to Zach, for some highly engaging discussions about academia but *not* about linguistics while in the field. And, to Jasmine and Maggie, who download an app every summer specifically so we'll still be able to talk while I'm in the field: I love you.

Thank you also to everyone in the W cubicles for their invaluable feedback. Additional thanks to my cohort for the many initial hypotheses and good-natured second opinions on my transcriptions. Laura merits a separate, special mention for coding the environment parser that accelerated my ability to hunt down minimal pairs and contrasting environments and gifted me the ability to spend my time elsewhere. Thanks also to everyone in the phonetics and phonology research group who provided feedback on the earliest drafts of this paper in its most chaotic form and to those in the documentation seminar for the time they have dedicated to discussing my preliminary analyses for the last three years. Endless thanks to Tony and Pattie for their careful feedback on each draft of this paper; this discussion would not be what it is without their insight. Thanks also to my students, who have gamely attempted to work through problem sets based on my data while reminding me of one of the reasons I entered graduate school at all.

And, finally, thanks to Sierra for her therapist and sister-approved strategies for coping with graduate school; your success is an inspiration. Thank you to my mother for talking me through my first year of teaching alongside everything else and motivating me

to suffer with her at the hands of our friend Jessica on YouTube, and to my father, for reminding me that I have interests outside of school and always being willing to talk about the Legion reboot, the Packers' good season, and the Red Wings' bad one when I need a break.

Consonants and Syllable Structure in Angaité (Enlhet-Enenlhet)

Paige Erin Wheeler, M.A.
The University of Texas at Austin, 2020

Supervisor: Patience Epps

This paper describes the consonant inventory and syllable structure in Angaité, a member of the Enlhet-Enenlhet family. Angaité utilizes 14 consonants which are described in this paper: five voiceless stops, three nasals, three fricatives, and three approximants. Two possible analyses of syllable structure are discussed, each of which results in slightly different restrictions on the distribution of glides within syllables as well as the relative frequency of different syllable structures. This paper also briefly sketches what is known of the vowel system in Angaité and how it relates to other suprasegmental features. Finally, this paper situates Angaité in the context of the language family and the broader Chaco region in which it is spoken, and it outlines key areas for future research in Angaité phonology.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Background

This paper describes the consonant inventory of Angaité (Enlhet-Enenlhet family, formerly Maskoy) and restrictions on consonant distribution within syllables and words. Angaité is highly endangered, with between 500 and 1,000 remaining speakers in the older generations (perhaps 60-80 years old) out of an ethnic population of around 4,000 (Fabre 2018; Gobierno Nacional 2012; Hammarström et al. 2019). The combined forces of Anglican missionaries and economic pressure toward non-indigenous employers have resulted in a community-wide shift to Paraguayan Guaraní, one of Paraguay's two official languages. The number of speakers is likely to decline along with the oldest generation in coming years; Angaité is not taught in schools, and it is no longer acquired by children.

Angaité is primarily spoken in the Presidente Hayes department in western Paraguay, one of three Paraguayan departments located in the South American Gran Chaco region which encompasses western Paraguay, northern Argentina, southern Bolivia, and a piece of western Brazil. Most Angaité people reside on reserves of indigenous land which are surrounded by privately-owned ranches (Fabre 2018). Some Angaité also live and work in the nearby Mennonite colonies or in Presidente Hayes, the departmental capital. This paper focuses specifically on Angaité as spoken in La Patria, a community located on a section of the Angaité's ancestral land that they successfully reclaimed from private companies approximately 30 years ago (Glauser 2019). La Patria is divided into 18 small villages, each called a *comunidad*, with slightly different linguistic identities depending on how much intermarriage has occurred with speakers of related languages. For example, the Carpincho *comunidad* is widely said to be 'pure' Angaité, while speakers in Comunidad 24 speak a more mixed variety. I focus here primarily on the Angaité spoken in La Leona, one of the largest and most central *comunidades*. The map in Figure 1.1 locates La Leona within the Presidente Hayes department.

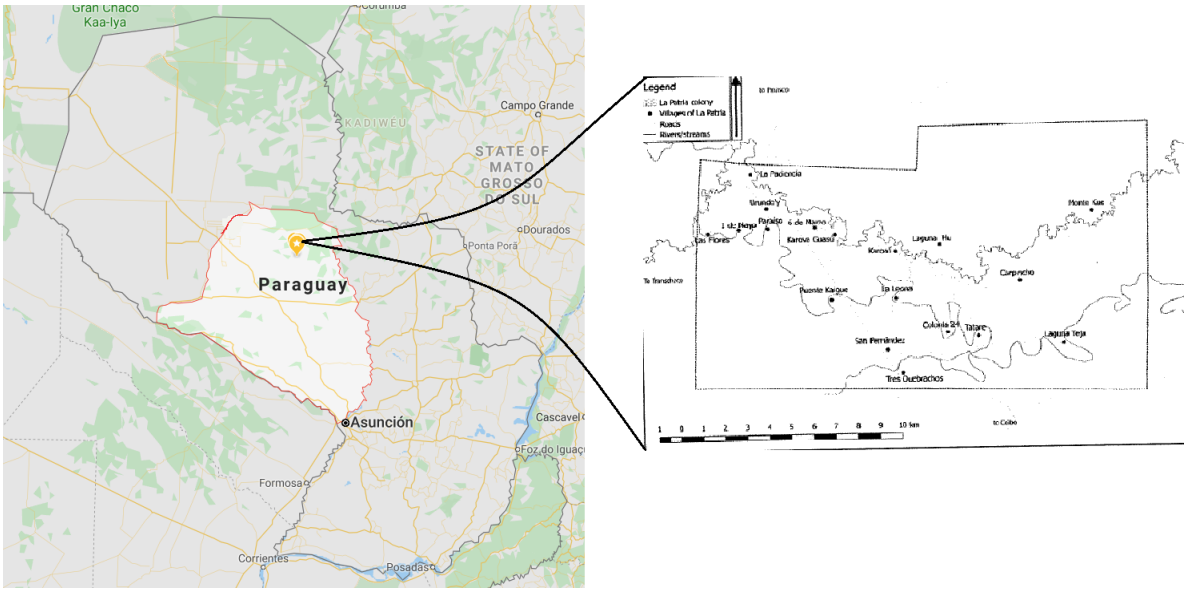

Figure 1.1: Map of La Leona and surrounding communities, inset from Glauser (2019)

Previous literature on Enlhet-Enenlhet languages is sparse, with the bulk of scholarship focused on anthropology (see D. Amarilla 2008; J. P. Amarilla 2006; Arenas 1981; Nobbs-Thiessen 2017; Unruh and Kalisch 1999a; Villagra Carrón 2014). In terms of linguistics, one of the earliest mentions of the family is Grubb (1911), an account by a missionary in the Chaco, which includes a brief language sketch as an appendix. Powys (1929) is a pedagogical grammar of Enlhet focused primarily on Biblical translation. Sušnik (1958; 1977) provides longer works which present a valuable snapshot of Enlhet during a period of rapid shift due to migrations and drastic economic and technological changes.

More recently, the working group *Nengvaanemkeskama Nempayvaam Enlhet*, led by Hannes Kalisch and Ernesto Unruh, has published brief descriptions of the family (Unruh and Kalisch 2003) and Enlhet specifically (Kalisch 2009).[1] Gomes (2009; 2013) provides a phonological and morphological sketch of Sanapaná, and two current U.S.-based PhD students are working on descriptions of Enxet (John Elliott, U. of Hawai'i) and Sanapaná (Jens van Gysel, U. of New Mexico). A dictionary project by Manolo Romero (*Nengvaanemkeskama Nempayvaam Enlhet*) and Raina Heaton (U. of Oklahoma) is also ongoing for Toba-Enenlhet.

[1]The bulk of Kalisch et al.'s work is focused descriptions of socioeconomic systems related to language discrimination in Paraguay, text collections, or pedagogical grammars, all of which are available at enlhet.org.

Even in the context of this minimally documented language family, Angaité is particularly understudied. The Paraguayan Secretaría de Políticas Lingüísticas (SPL) is in the process of producing a collection of videos in Angaité with subtitles in Guaraní, but this project will not attempt comprehensive language documentation or linguistic description.[2]

This paper describes the consonant contrasts in Angaité and phonotactic restrictions on their distribution within syllables and words. This description is intentionally narrow in scope. Angaité's vowel system is linked to a variety of prosodic features including stress, pitch, phonation type, and segment length. In turn, these suprasegmental features are affected by morphological processes like affixation, which is particularly evident in verbs due to the rich morphology in verbal constructions. Therefore, some level of prosodic and morphological analysis is necessary to approximate a reasonably thorough description of Angaité vowels; the corpus developed in my pilot study is not suited to this task. However, Angaité's phonemic consonants have comparatively fewer allophones and are less implicated in suprasegmental processes. Consequently, I have opted to provide a relatively precise description of consonant contrasts and phonotactics based on my current corpus. This work highlights directions for future research while grounding subsequent investigation in a well-motivated analysis of a subset of the phonological system. Consonants are of course not exempt from interaction with prosodic and morphological processes, and I anticipate that the current work will be refined over time. Nevertheless, I do not anticipate that future research will reveal facts so unanticipated as to make the current description obsolete.

Even with this narrow scope, the current work represents a significant contribution to both the description of one particular endangered and under-documented language and to the larger project of illuminating patterns of language contact, change, and small-scale multilingualism among the diverse groups residing in the Gran Chaco. Previous work on Enlhet-Enenlhet languages has not undertaken a phonological description of Angaité. Furthermore, Gomes (2013) focuses on Sanapaná phonology and morphology, and describing Angaité opens the door to comparison between the two Southeastern Enlhet-Enenlhet languages. Comparison within sub-branches of the family is a first step toward historical work

[2]The SPL publishes their work on their YouTube channel, here: https://www.youtube.com/channel/UCXjY_IAv-IacFu9T-PCSRqA/videos.

both within the Enlhet-Enenlhet family and among the six language families (28 languages) and two isolates that are spoken in the Chaco (Durante 2011). Historical work on Chaco languages in general has long been the missing piece in discussions about dynamics of language contact in the region, and Enlhet-Enenlhet languages have been often overlooked due to lack of data and description. This work offers the possibility to include Angaité in these broader discussions.

Recent phonological sketches of Sanapaná suggest the following consonant inventory for that language, which I take as a starting point for describing Angaité. van Gysel (2017) revises Gomes' phonological sketch of Sanapaná, arguing for 13 consonant contrasts, as in Table 1.1.

Manner \ Place	Bilabial	Alveolar	Palatal	Velar	Glottal
Plosive	p	t		k	ʔ
Nasal	m	n		ŋ	
Fricative		s			h
Lateral Fricative		ɬ			
Approximant		l	j	w	

Table 1.1: Sanapaná consonants, from van Gysel (2017)

Unruh and Kalisch (2003) also state that Enlhet-Enenlhet languages except Toba-Enenlhet contrast vowels based on length. van Gysel (2017) notes that while some words in Sanapaná are consistently pronounced with long vowels and others with short ones, his data do not include any minimal pairs that contrast on vowel length and therefore he does not analyze it as phonemic. This paper, though not primarily concerned with vowel length, does briefly discuss it, since previous sources conflict on the importance of length distinctions.

Glottalized sonorants and glottal stops are of particular interest in this discussion. Summaries of regional features (see Campbell 2012, 2013) suggest that Enlhet-Enenlhet languages utilize glottalized sonorants (nasals and vowels), though the distribution and phonemic status of these sounds is not clear. Additionally, Gomes (2013) describes processes of consonant epenthesis, particularly of a glottal stop, due to apparent restrictions on word final open syllables. My own data on Angaité show creaky voicing and pitch changes in

vowels that may correspond to underlying /ʔ/. A second goal for this analysis is to determine if Angaité participates in the patterns that have been described in related languages or whether glottalized consonants and glottal stops behave differently in Angaité.

In the following sections, I describe the documentary corpus used here, including methods, consultants, and recording equipment (§1.1); outline the Enlhet-Enenlhet practical orthography used in this paper (§1.2); and provide a brief typological sketch of Enlhet-Enenlhet languages based on previous research as context for the current phonological analysis (§1.3). The bulk of this paper is concerned with detailing syllable structure and consonant contrasts. Chapter 2 details the consonant contrasts within each natural class, returning to more specific patterns within syllables or words where relevant. With the consonant contrasts established, Chapter 3 discusses syllable structure and the distribution of consonants within syllables. Finally, Chapter 4 briefly addresses what can currently be said about Angaité vowels, and Chapter 5 sums up the current understanding of Angaité's phonological inventory and directions for future research.

1.1 Practical Orthography

The practical orthography used in this paper is adapted from the one developed by *Nengvaanemkeskama Nempayvaam Enlhet* for Enlhet (see Unruh and Kalisch 1999b) and is largely the same as the orthography used to transcribe Angaité in other published text collections. It represents only the phonemes of the language, based on my current analysis of the phoneme inventory.[3] The alphabet used to transcribe Angaité is listed in Table 1.2.

This alphabet uses a 1:1 correspondence between orthography and the IPA. Where the alphabet differs from the IPA representation, the IPA symbol appears in brackets next to the letter that represents it in this chart. Uvular stop /q/ is a slightly special case which will be discussed in further detail in §2.1.1

Other Enlhet-Enenlhet languages also use <g> (phonetic characterization unclear).

[3]This orthography may evolve based on community consensus and will likely eventually represent some underlying forms which differ from surface realizations, but this paper does not attempt spelling standardization.

Manner of Articulation	Labial	Alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p	t		k	q	’ [ʔ]
Nasal	m	n		ng [ŋ]		
Fricative		s				h
Lateral Fricative		lh [ɬ]				
Approximant		l	y [j]	v [w]		

Table 1.2: Angaité practical orthography

Unruh et al. (2003) define the letters in Toba-Enenlhet based on Guaraní or Spanish, with the exception of <lh>, <v>, <g>, and <q>, for which they present example words in Toba-Enenlhet. Letters <lh>, <v>, and <q> represent the same sounds in Toba-Enenlhet as in this Angaité orthography. Other sources, including Unruh and Kalisch (1999b) about Enlhet orthography, however, do not use <g>, leaving me somewhat unsure as to what it represents. Angaité's inventory does not include a voiced velar stop or similar sound, and I do not use this letter in the Angaité practical orthography. Unruh et al. (2003) also use <i> to represent the palatal glide /j/, rather than <y>, which is used in the orthographies for Enlhet and Guaná. Since <y> is in wider use than <i>, I have opted to use it here in the interest of reducing orthographic variation.

The three phonemic vowels in Angaité /a, e, o/ are written with the same graphs in the orthography. Long vowels are represented with a double letter. An additional reason that I have avoided using <i> to represent /j/ is to avoid confusion with the vowel [i], which is (as currently analyzed) allophonic in Angaité and not represented in the orthography. In contrast, J. P. Amarilla (2006) and Domaniczky and Imaz (2006) use three vowels to transcribe Angaité, <a, i, o>. Neither work defines their orthography, but <i> here clearly represents a vowel and not the palatal glide. Though this system has been used in previous publications specific to Angaité, I avoid using <i> to represent a vowel because, as far as I am aware, no other Enlhet-Enenlhet language uses <i> for a vowel phoneme and because representing a phonemic mid vowel (/e/) as <i> is needlessly complicated.

In this paper, the first *italicized* line of examples is written in Angaité practical orthography, with the second line a phonemic transcription in the IPA. Where necessary to demonstrate a phonological process, the second line of a gloss may be a phonetic transcrip-

tion, in which case it is [bracketed]. Examples whose first lines are blue are clickable for audio. Long vowels appear in both the orthographic and phonemic lines when the vowel is notably longer than other vowels in the word; vowel length may turn out not to be phonemic, but since it is not fully analyzed in this paper I have transcribed long vowels to avoid eliding potentially relevant detail. Phonation type, pitch, and stress are not indicated in these transcriptions unless specifically noted.

1.2 Methodology

This paper is based on data collected during my fieldwork in La Patria in July 2019 in collaboration with speakers from Karoa'i, Comunidad 24, Koralón, and La Leona. My primary elicitation collaborators were Eusebelina González and Eusebia Fernández, with transcription and translation of texts done with help from Damasio Flores and Amancia Samalleco. Data used in this analysis is archived at AILLA in Wheeler 2019–. Data was audio and video recorded using a Zoom Q8 recorder attached to an external omnidirectional condenser microphone (Sony ECM-MS907). The recorder created both an integrated audio-video .MOV track and a separate (22/44.1Hz stereo) .WAV track. Recordings were made in semi-structured elicitation sessions with two Angaité speakers from La Leona: Eusebelina González (EG) and Eusebia Fernández (EF). Each session lasted approximately an hour with 18 total hours of elicited data recorded. During these sessions, the microphone was placed between the speakers and set with a 120° range in order to pick up both speakers as well as my own voice. The Zoom Q8 internal omni-directional mics were not used in an attempt to limit the interference of background noise on the recordings.

Elicitation sessions were structured to address two main goals: investigate the details of an omnipredicative analysis in Angaité and collect a word-list for phonological analysis. The first line of investigation required eliciting simple sentences, including possessive noun phrases, copulas, negation, and temporal adverbs. This task involved both basic translation from Guaraní ("how do you say '____' in Angaité?") and scenarios established with the use of pictures. Pictures used to prompt sentences included: a chicken, a dog, men and women in both groups and alone, pictures of family members, and people looking at various objects (e.g.

man looking at a chicken, woman looking directly at camera, etc.). Sentences for this task were constructed first in Spanish and then translated to Guaraní to be used as translation prompts, or constructed by me in Angaité and backtranslated by EG and EF into Guaraní. The latter method was only used to check translations or complete verbal paradigms with forms that were difficult to elicit with translation from Guaraní (e.g. masculine or plural forms of verbs).

The second task was accomplished via translation of a word list drawn from a variety of sources: 158 items were translated from a list of Toba-Enenlhet words collected during a pilot trip in July 2018, 66 items came from the Key and Comrie (2015) Intercontinental Dictionary Series (IDS) list for Lengua (Enlhet), 190 terms came from the Sanapaná IDS list, and approximately 40 plant and animal terms were translated from a list of Guaraní names for Chaco-specific flora and fauna developed in 2018. A few elicited tokens were also prompted by showing pictures of plants, animals, or insects taken in La Patria.[4] See a complete word list of elicitation prompts (and their sources) in Appendix C.

These two tasks were not performed independently. We moved between tasks within and between sessions as EG and EF tired of repeatedly performing one task. EG and EF also volunteered sentences based on the elicitation prompts, either offering full sentences with an elicited word in context or elaborating prompts with related sentences. Some portion of each day's work was also dedicated to re-checking information from previous sessions, resulting in multiple tokens of many words produced on different days and by more than one speaker. Because of the scarcity of data on Angaité, this analysis draws on all these types of data and is not limited to words produced entirely in isolation.

Unless noted as stemming from a source aside from my own fieldwork, the original recordings from which the Angaité examples in this paper are drawn are available in Wheeler (2019–). Parenthetical codes that appear with each token indicate the date (YYYYMMDD), session number (1, 2, or 3), and item number within a session of that token. The archive deposit is organized by date, with bundles and files labeled by the same date that appears

[4]The list of Toba-Enenlhet words from 2018 was also elicited at that time based on Spanish translations of the IDS lists for Enlhet and Sanapaná. I list it separately here because EG and EF translated specifically from Toba-Enenlhet prompts and not directly from the IDS list.

in the data citation. Each file in the archive also has a unique, 4 digit recording number. A data citation which refers to session 1 corresponds to the file name with the lower recording number. For example, (20190708, 3, 10) refers to the 10th item in the third elicitation session on July 8, 2019. The July 8 recording bundle includes files with recording numbers 0004, 0005, and 0006. The third elicitation session corresponds to recording 0006. When relevant, speakers are cited in these parenthetical citations by initials. Some data are also drawn from my 2018 fieldwork on Toba-Enenlhet working with Manolo Romero in Nueva León (Departamento Boquerón); these recordings are not (yet) archived and therefore not indexed to a publicly available source. Data from sources other than my own fieldwork are cited by author last name, year, and page number on which the original example appears and have not been altered unless otherwise noted.

1.3 Language Background

Though a grammar sketch of Angaité is beyond the scope of this paper, this section provides a brief typological sketch of the family in order to give an idea of the shape of Enlhet-Enenlhet morphosyntax, apparent morphophonological processes, and the categories that are anticipated to surface in future work.

1.3.1 Family classification

Angaité is one of six Enlhet-Enenlhet languages. Unruh and Kalisch (2003) establish a classification of the family based on synchronic similarities in lexicon, grammatical structures, and phonemic inventories which divides the family into two main groups, Eastern and Western, with the Eastern group subdivided into Northern and Southern. The Western group contains Enlhet and Enxet (often called Lengua Norte and Lengua Sur, respectively). The Southeastern group comprises Angaité and Sanapaná, and Guaná and Toba-Enenlhet form the Northeastern group. Figure 1.2, adapted from Unruh and Kalisch (2003) indicates the main locations of Enlhet-Enenlhet groups. The western border on the map inset is the border between Paraguay and Argentina, and the eastern border is formed by the Paraguay River to the South and the Brazilian border to the north.

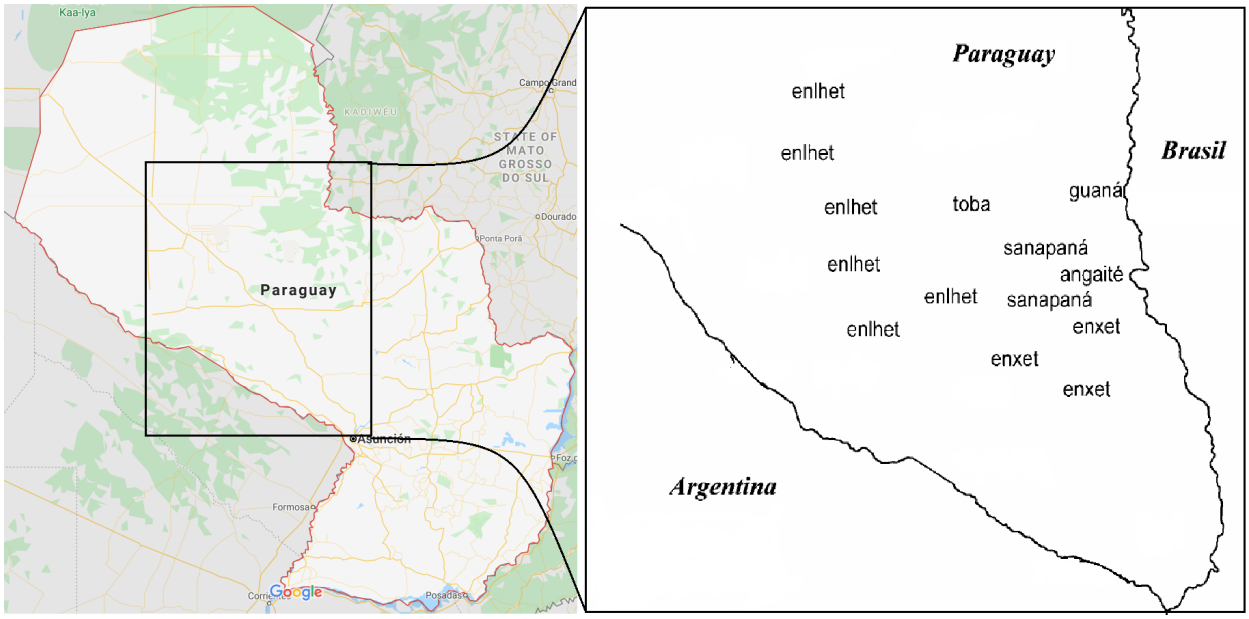

Figure 1.2: Location of Enlhet-Enenlhet groups in the Paraguayan Bajo Chaco, from Unruh and Kalisch (2003)

However, Kalisch and Unruh stress the diversity of dialects within each language and therefore distinguish them primarily based on speakers' ethnic self-identification. They also argue that geographic location is highly influential in determining which features appear in which variety, with each basically mutually intelligible with its geographic neighbors. Because of this high dialectal diversity, they argue that the family is better understood as a continuum of dialects with six 'nuclei' rather than discrete languages.

This classification so far lacks description of each language that could be leveraged for diachronic analysis to identify shared innovations and, if present, distinct subgroupings. Detailed descriptive work can also contribute to the goal of determining which features have been shared via contact between Enlhet-Enenlhet languages in order to better ground an analysis of the family as a dialect continuum formed by continuous interaction between groups. Where relevant, this paper compares my current analysis of Angaité to Gomes (2013), which is primarily a phonological description of Sanapaná. Differences between these two very closely related languages—e.g. which segments (if any) are epenthetic, the interaction between morphology and syllabification, stress assignment rules—suggest fruitful directions for future research, particularly with respect to the interaction between phonology

and morphology.

1.3.2 Areal phonological features

Existing descriptions of Enlhet-Enenlhet languages suggest that they include some phonological features that have also been proposed as Chaco areal features. See Table 1.3, which compiles information about Enlhet-Enenlhet languages from Gomes (2013), Powys (1929), van Gysel (2017), Unruh and Kalisch (2003), Unruh et al. (2003), and my own fieldwork. The list of proposed phonological areal features is drawn from Campbell (2012; 2013), Campbell and Grondona (2010), and González (2014; 2015). Cells are marked with '?' if sources disagree about the presence or absence of a feature or if a blanket statement is made about the entire family based on evidence from only a subset of its members. The large number of question marks is striking; gaps and contradictions in existing descriptions raise a variety of issues to be addressed by future descriptive projects. Furthermore, scholars interested in areal features agree that historical work grounded in robust descriptions is necessary to flesh out the patterns observed in some Chaco languages and determine the extent to which each language participates in them.

Feature	**Enlhet**	**Enxet**	**Sanapaná**	**Toba**	**Angaité**	**Guaná**
Ejective sounds	No					
Glottalized sonorants	Yes?					
Phoneme /ɬ/	Yes					
Lateral v. liquid based on place/manner	Yes					
3-4 vowel systems	Yes					
Voiceless nasals	N?	N?	N?	N?	N	N?
Vowel harmony	Y?	Y?	Y	Y?	Y?	Y?
No voiced stops	Y?	–	Y	Y?	Y	–
Uvular sounds	N (some)	Y	N	Y	Y	Y

Table 1.3: Presence of proposed Chaco phonological features in Enlhet-Enenlhet languages

1.3.3 Enlhet-Enenlhet morphosyntax

Outside the realm of phonology, Enlhet-Enenlhet languages are agglutinating, marking event participants, associated motion, and tense, aspect, and mood (T/A/M) categories within the verbal construction. As a result, the order of sentence constituents is generally

flexible and arguments are often not independently expressed. When both arguments are overtly expressed as full NPs, the least marked order is VOA (see van Gysel (2017) for Sanapaná). For example, see (1), (2), and (3) from Angaité.

(1)

'aktooma'	*kelasmatko'ok*	*haapo'*
ʔak-tooma**ʔ**	kelasmat-ko**ʔ**ok	haapo**ʔ**
2/3F.DIR-eat	fish-DIM	bird
V	O	A

'The bird eats the fish' (20190722, 2, 27)

(2)

'apyengkek	*yemen*	*'apkenao'*
ʔap-je**ŋ**kek	jemen	ʔap-kenao**ʔ**
2/3M.DIR-drink	water	2/3M-man
V	O	A

'The man drinks water' (20190717, 1, 34)

(3)

'akvetay'a	*'angkelvana*	*'apkenao'*
ʔak-wetaj**ʔ**a	ʔa**ŋ**-kelwana	ʔap-kenao**ʔ**
2/3M.DIR-see	2/3F-woman	2/3M-man
V	O	A

'The man sees the woman' (20190717, 1, 10.5)

Animacy also plays into word order in Angaité. For example, when an animate A is paired with an inanimate O argument, the order is VAO. For example, see (4):

(4)

'apayvaskama'	*'apkenao'*	*'aphaspong*
ʔap-ajwaskama**ʔ**	ʔap-kenao**ʔ**	ʔap-haspo**ŋ**
2/3M.DIR-play	2/3M-man	2/3M.POS-guitar
V	A	O

'The man plays guitar' (20190717, 2, 42)

However, clearly animacy is not the only factor that influences word order, as VOA order still surfaces in cases with an animate A and inanimate O, as in (2). I suspect that information structure, alongside animacy, plays a role in determining whether order is VOA or VAO. Kalisch (2019) argues for a similar role of information structure in clausal order in Enlhet, with cognitive accessibility governing both order and which elements can be elided.

As in many Chaco languages, most categories associated with verbs appear as suffixes or enclitics. Previous literature indicates that Enlhet-Enenlhet distinguishes three tenses as well as a variety of modal and aspectual categories (Powys 1929; Susnik 1977; van Gysel 2017). These T/A/M categories are expressed either as suffixes or enclitics; van Gysel (2017) describes both tense and aspect suffixes as well as three temporal discourse particles that function more like enclitics. Enlhet-Enenlhet languages also use a well-differentiated system of associated motion markers that include information about direction and path, as well as other more nuanced categories. Languages in this family also include some marking of evidentiality. These categories are also marked as suffixes or verb stem alternations.[5]

Enlhet-Enenlhet languages mark gender (masculine or feminine), number (singular or plural) and person (first or non-first) of the verbal arguments with portmanteau markers that attach before the verb stem. Typically, these markers refer to either the A or S argument. First-person distinguishes singular and plural, and non-first person marks the gender of the argument. The languages also utilize an inverse marking system. One set of markers indicates number of first-person A arguments (Kalisch (2009) calls them 'central arguments') and gender of non-first person A arguments when the O ('less prominent argument') is also non-first person. A different set appears when A is a non-first person acting on a first-person O; in this case, the marker encodes the number of the O argument, not the gender of the A. Kalisch (2009) divides these markers into three paradigms, with Paradigm I used for what seems to be unrealized events, Paradigm III marking subordinate clauses, and Paradigm II used elsewhere. See Table 1.4, which presents the markers as shown in Kalisch (2009).[6]

[5]Sušnik (1977) includes categories such as whether or not the person takes a direct or curving path, whether they will stay at the destination for a long period of time or visit and then leave again, and whether the journey includes multiple stops.

[6]Kalisch does not propose underlying forms of these morphemes or describe the environments in which

Marker	Paradigm I	Paradigm II	Paradigm III
$1SG_{DIR}$	*a-* ∼ *angv-*	*ak-* ∼ *ay-*	*sek-* ∼ *sey-* ∼ *s-*
$1PL_{DIR}$	*ang-* ∼ *am-* ∼ *an-*	*neng-* ∼ *nen-* ∼ *nem-*	
$2/3F_{DIR}$	*ka-* ∼ *ngka-*	*ang-* ∼ *an-* ∼ *am-* ∼ *angk-*	*ak-* ∼ *ay-*
$2/3M_{DIR}$	*e-* ∼ *engy-*	*ap-* ∼ *apk-*	
$1SG_{INV}$	*he-* ∼ *hey-*	*e-* ∼ *ey-*	*se-* ∼ *sey-*
$1PL_{INV}$	*heng-* ∼ *hen-* ∼ *hem-*	*eng-* ∼ *en-* ∼ *em-*	*seng-* ∼ *sen-* ∼ *sem-*

Table 1.4: Person markers in Enlhet, from Kalisch (2009)

Gender is a relevant category in Enlhet-Enenlhet even outside of verbal constructions. Nouns in Enlhet-Enenlhet languages have gender, which is expressed for un-possessed nouns via either a person marker on the verb or on the noun itself. According to Kalisch (2009), the gender markers on nouns are often optional, but even with no overt marker of gender, speakers are aware of a noun's gender, suggesting that it is a relatively salient aspect of nouns. For example, see the data from Toba-Enenlhet in Table 1.5.

Masculine			**Feminine**		
pelhasep	[peɬasep]	'night'	*paga*	[paga]	'mosquito'
tenoq	[tenoq]	'cat'	*posek*	/posek/	'parakeet'
songkong	[soŋkoŋ]	'tuca tuca'	*konemaqtek*	[konemaqtek]	'dove'
ka'a	[kaʔa]	'tereré' (loan from Guaraní)	*sepop*	[sepop]	'3-banded armadillo'
melhma	[meɬma]	'coals'	*pehe*	[pehe]	'lizard'
popeye'	[popejeʔ]	'deer'	*eyam*	[ejam]	'south wind'
kelasma'	[kelasma]	'fish'	*sepo'*	[sepo]	'manioc'
meyva'	[mejwa]	'puma'			

Table 1.5: Noun gender in Toba-Enenlhet, from fieldwork July 2018

The gender of the noun seems to be arbitrary, with both masculine and feminine categories including mammals, reptiles, birds, fishes, plants, tools, foods, and weather. Feminine gender is the default in both Angaité and Toba-Enenlhet, indicated by loan words, which are assigned feminine gender. For example, see (5) and (6), which both have feminine gender.

each allomorph appears. Based on my Angaité data, I hypothesize that the allomorph that appears on the left in each cell is the underlying form, with other allomorphs phonologically conditioned by the stem to which they attach.

(5) From Toba-Enenlhet

ka'a
kaʔa
'yerba mate' (loan from Guaraní)

(6) From Angaité

vatka'	*'ayempehek*
watkaʔ	ʔa-jempehek
cow	2/3F.POS-skin
'leather'	(20190722, 2, 66)

Possessed nouns receive a marker that indicates the gender of second/third-person possessors, rather than the possessum. For example, see (6), where *'a-* corresponds to the feminine gender of *watka'*, 'cow'. At least in Enlhet, for which they are briefly described in Kalisch (2009), these markers are similar, though not identical to, the Paradigm II masculine, feminine, and inverse verbal argument markers.

Because the person markers appear on many unpossessed nouns and all possessums, Kalisch (2009) argues that Enlhet-Enenlhet languages are omnipredicative, meaning that all nouns have an underlying predicative function identical to verbs and their default use is as the ground to a proposition. Words fall into a single class of predicates and must be derived into (more marked) argument functions. On this analysis, person markers indicate argument saturation, both for verbs—where they denote event participants—and for nouns, with markers indexing either a possessor or a real-world referent. However, per Kalisch, these markers can often be omitted and in fact do not appear in many of his examples. Other literature on Enlhet-Enenlhet suggests that this analysis may not be the most fitting option. Sušnik (1977) divides nouns in Enlhet into three categories: obligatorily possessed nouns, nominalized verbal constructions, and all other nouns (mostly nouns referring to concrete things in the natural world). This final category actually cannot appear in possessive constructions or ever take the markers on which Kalisch's argument hinges. This paper leaves the issue

of omnipredicativity open, but preliminary fieldwork indicates that, at least in Angaité, the question of the function and behavior of the person markers is not as straightforward as a strict omnipredicative analysis predicts.

Like in the realm of phonology, Enlhet-Enenlhet languages also participate in Chaco patterns of morphosyntax, as much as they can be said to exist. AVO order, prefixed verbal person markers, active-stative alignment, nominal tense, associated motion markers, grammatical gender marked on demonstratives, plural object suffixes, genitive classifiers for domestic animals, and a clusivity distinction in first-person plural have been proposed as Chaco areal features (see sources discussed above for regional phonological features). Enlhet-Enenlhet languages also participate in many of these patterns, as indicated in Table 1.6.

Feature	**Enlhet**	**Enxet**	**Sanapaná**	**Toba**	**Angaité**	**Guaná**
SVO	No					
Nominal tense	No?					
Active-stative	Yes?					
Person prefixes	Yes					
Associated motion	Yes					
Gender marking on demonstratives	Yes					
Genitive classifiers for domesticated animals	Yes					
Plural object suffixes	Y	Y?	Y?	Y?	Y?	Y?
1PL clusivity	N?	N?	N	N	N	N?

Table 1.6: Structural features shared in Enlhet-Enenlhet

As Table 1.6 summarizes, Enlhet-Enenlhet languages use large associated motion systems, mark grammatical gender on demonstratives, include genitive classifiers for domestic animals, have preverbal verbal person markers, and may also include plural object suffixes and active-stative alignment systems similar to other Chaco languages. However, like the overview of phonological features, this summary highlights substantial gaps for future description.

Chapter 2

Angaité Consonant Contrasts

Consonants in Angaité pattern together primarily based on manner of articulation. This section takes each natural class in turn. Sections 2.1 through 2.5 discuss: plosives, nasals, fricatives, laterals, and glides. When relevant, patterns that attend to place of articulation, rather than manner, are discussed within each section. Consonants are shown to contrast word-initially; word-finally; and word-medially, as both syllable onsets and codas.

For the purposes of this discussion, the maximal syllable is CVCC. Since all words in the Angaité corpus begin with consonants, I assume that Angaité, like many languages, prefers syllables with onset consonants. I placed a syllable boundary before every (C)V sequence, so a CVCV sequence is syllabified CV.CV and a VCCV sequence, VC.CV. Headless syllables appear only in CVVC sequences, syllabified as CV.VC. In these data, /j/ is the only phoneme which can appear in a complex coda, and only before /k/, so the maximal syllable is /CVjk/.

2.1 Plosives

This section discusses Angaité's five voiceless stops: /p, t, k, q, ʔ/. After establishing the basic contrasts of /p, t, k, ʔ/, I discuss some alternations which affect these phonemes. Section 2.1.1 considers /q/ specifically, since it is a somewhat special case. Section 2.1.2 discusses the alternation between /ʔ/, /h/, and creaky voice, §2.1.3 discusses a /t/∼ /ʔ/ ∼ /k/ alternation at morpheme boundaries, and §2.1.4 describes a potentially idiosyncratic /p/ ∼ /w/ alternation.

The /p, t, k, ʔ/ phonemes appear word-initially, word-medially, and word-finally. /p, t, k/ appear often as both onsets and codas word-medially, while /ʔ/ appears primarily in onset position word-medially with only a few examples of medial coda /ʔ/.

The elicited data includes several (near) minimal pairs that demonstrate plosive contrasts, presented in Table 2.1. Contrasting segments are **bolded**.

Contrasting *p* and *t*					
***t**eyam 'eten* /tejam ʔeten/	'fog'	(20190722, 1, 21)	***p**eeyam* /peejam/	'honey'	(20190723, 2, 28)
*'ak**t**ong* /ʔaktoŋ/	'your/her arm'	(20190715, 1, 37)	*'ak**p**ong* /ʔakpoŋ/	'shell'	(20190715, 1, 37)
*'a**t**ava'* /ʔatawaʔ/	'your/her husband'	(20190722, 1, 31)	*'a**p**ava'* /ʔapawaʔ/	'tagua'	(20190722, 1, 46)
Contrasting *k* and *p*					
***k**aana'* /kaanaʔ/	'type of bird'	(20190717, 2, 28)	***p**aana'* /paanaʔ/	'heron'	(20190710, 2, 6)
*'a**k**toma'* /ʔaktomaʔ/	'You/she eats'	(20190708, 3, 32)	*'a**p**toma'* /ʔaptomaʔ/	'You/he eats'	(20190708, 3, 59)
*'a**p**taykamaha'* /ʔaptajkamahaʔ/	'You/he works'	(20190716, 1, 44)	*'a**k**taykamaha'* /ʔaktajkamahaʔ/	'You/she works'	(20190722, 1, 46)
Contrasting *p* and *'*					
*lheya**p*** /ɬeap/	2/3M	(20190708, 3, 59)	*lheya**'*** /ɬeaʔ/	2/3F	(20190708, 3, 8)
Contrasting *k* and *'*					
*ko'o**'*** /koʔoʔ/	1SG	(20190722, 1, 6)	*-ko'o**k*** /-koʔok/	DIM.PL	(20190722, 1, 29)

Table 2.1: (Near) minimal pairs demonstrating plosive contrasts

Along with the minimal pairs in Table 2.1, the data provide robust evidence for contrasting environments. For example, these four plosives appear word-initially before all three phonemic vowels. Because Angaité does not allow onset consonant clusters, the plosives always appear before a vowel word-initially. Examples of word-initial environments appear in Table 2.2.

		Word-initial /p/	
paana'	/paanaʔ/	'heron'	(20190710, 2, 6)
panaktema'	/panaktemaʔ/	'medicine'	(20190708, 2, 18)
pelhapen	/peɬapen/	'rhea'	(20190708, 1, 36)
peletav	/peletaw/	'knife'	(20190710, 2, 52)
poktem	/poktem/	'timbo'	(20190718, 13)
popen	/popen/	'guinea hen'	(20190717, 2, 61)
		Word-initial /t/	
taalha'	/taaɬaʔ/	'fire'	(20190708, 1, 1)
taama'	/taamaʔ/	'thread, cord, lasso'	(20190710, 2, 54)
teeves	/teewes/	'algarrobo negro'	(20190708, 2, 4)
tevovok	/tewowok/	'great horned owl'	(20190717, 2, 29)
topole'	/topoleʔ/	'ant'	(20190710, 2, 16)
tomahang	/tomahaŋ/	'rabbit'	(20190723, 1, 41)
		Word-initial /k/	
kaava'	/kaawaʔ/	'liana, Dutchman's pipe'	(20190717, 2, 34)
kalenmaqtek	/kalenmaqtek/	'dove'	(20910708, 1, 63)
kenavet	/kenawet/	'wasp'	(20190722, 1, 67)
kelhvoye'	/keɬwojeʔ/	'soon, fast'	(20190723, 2, 11)
kelyaqhava'a	/keljaqhawaʔa/	'collared peccary'	(20190708, 1, 50)
kelvanema'et	/kelwanemaʔet/	'wolf fish'	(20190708, 1, 45)
		Word-initial /ʔ/	
'apketka'	/ʔapketkaʔ/	'your/his child'	(20190708, 3, 29)
'askanmah	/ʔaskanmah/	'I sow'	(2019708, 2, 28)
'enyaalheng	/ʔenjaaɬeŋ/	'my older sister'	(20190722, 1, 5)
'engva'	/ʔeŋwaʔ/	'my/our hair'	(20190722, 1, 13)
'onyenek	/ʔonjenek/	'I drink (tereré)'	(20190708, 1, 17)
'olhtenaksek	/ʔoɬtenaksek/	'I tell'	(20190716, 1, 43)

Table 2.2: Word-initial /p, t, k, ʔ/

These four plosives also appear word-finally, as in Table 2.3.

		Word-final /p/	
hatsap	/hatsap/	'toucan'	(20190718, 1)
tahap	/tahap/	'ashes'	(20190708, 1, 25)
pelhasep	/pełasep/	'night'	(20190708, 3, 12)
'anyep	/ʔanjep/	'field'	(20190722, 1, 2)
mepop	/mepop/	'skunk (Sp. zorrillo)'	(20190708, 1, 64)
lhepop	/ɬepop/	'dust, dirt, earth'	(20190722, 1, 50)
		Word-final /t/	
pa'at	/paʔat/	'grass'	(20190717, 2, 12)
kelyamenanat	/keljamenanat/	'Azara's night monkey'	(20190717, 2, 12)
leklakmet	/leklakmet/	'thank you'	(20190710, 2, 85)
yetamayet	/jetamajet/	'anaconda'	(2019017, 2, 70)
		Word-final /k/	
vatkahak	/watkahak/	'paper, notebook'	(20190716, 1, 61)
yaamet 'awhak	/jaamet ʔawhak/	'tree root'	(20190723, 1, 22)
'aypehek	/ʔajpehek/	'fish scale' (F)	(20190710, 2, 14)
yaasek	/jaasek/	'salt'	(20190710, 3, 25)
'ahankok	/ʔahankok/	'her thing/pet'	(20190716, 1, 26)
yamav'aswok	/jamawʔaswok/	'small lagoon'	(20190723, 1, 5)
		Word-final /ʔ/	
vaapa'	/waapaʔ/	'rat'	(20190708, 2, 93)
yalhpa'	/jaɬpaʔ/	'mud'	(20190708, 1, 18)
latsehe'	/latsehеʔ/	'corn'	(20190720, 1, 31)
'astengke'	/ʔasteŋkeʔ/	'I sleep'	(20190720, 1, 46)
'anyalho'	/ʔanjałoʔ/	'(woman's) sister'	(20190719, 1, 9)
sepo'	/sepoʔ/	'manioc'	(2019020, 1, 29)

Table 2.3: Word-final /p, t, k, ʔ/

My data do not include examples of word-final /t/ following /o/. Given that /t/ can appear adjacent to /o/ in other positions and that the other plosives can appear word-finally, I assume that this is an incidental gap rather than a restriction on word-final /ot/ sequences. Further evidence for this assumption is provided by the fact that word-internally /t/ does appear in coda-position following /o/ as in (7).

(7) *ne.nga.vot.ma'* *sosenhe'*
neŋ-awotmaʔ sosenheʔ
1PL.DIR-make net.bag
'We make net bags' (20190722, 2, 71)

In general /o/ appears less frequently than /a/ and /e/ in Angaité. Again, this fact combined with the behavior of /t/ in word-medial codas predicts word-final *ot* sequences in a larger corpus.

The phonemes /p, t, k, ʔ/ also appear word-medially as both syllable onsets and codas, though /ʔ/ is not attested as a word-medial coda. For the other three, onset position is more common than coda; codas in general are less frequent than onsets in Angaité. The examples in Table 2.4 show plosives in onset position and Table 2.5 shows the plosives in word-medial coda position. Syllable boundaries are marked in the orthography.

Word-medial onset /p/			
*ya.me.yek.**p**a'*	/jamejekpaʔ/	'tucatuca'	(20190717, 2, 2)
*yaa.**p**a'*	/jaapaʔ/	'money'	(20190723, 1, 59)
*'ak.mo.**p**ay.'a'*	/ʔakmopajʔaʔ/	'white (F)'	(2019710, 3, 60)
*'e.yem.**p**e.hek*	/ʔejempehek/	'my/our skin'	(2019720, 2, 44)
*'ap.ma.lhem.**p**e.nek*	/ʔapmaɬempenek/	'your/his cheek'	(20190715, 1, 45)
*pe.lha.**p**en*	/peɬapen/	'rhea'	(20190722, 1, 65)
*lhe.**p**op*	/ɬepop/	'dust'	(20190722, 1, 50)
*me.**p**op*	/mepop/	'bat'	(20190708, 1, 64)
*kas.**p**o.ma'*	/kaspomaʔ/	'cigarette'	(20190710, 2, 61)
Word-medial onset /t/			
*'ak.**t**ay.ka.ma.ha*	/ʔaktajkamaha/	'You/she works'	(20190720, 2, 14)
*nen.**t**alh.ne.ma'*	/nentaɬnemaʔ/	'our clothes'	(20190717, 1, 51)
*see.**t**a'*	/seetaʔ/	'grandfather'	(20190708, 2, 23)
*pa.nak.**t**e.ma'*	/panaktemaʔ/	'medicine'	(20190708, 2, 18)
*yaq.**t**e.pa'*	/jaktepaʔ/	'pumpkin'	(20190708, 2, 28.5)
*'aq.**t**ek*	/ʔaqtek/	'your/her eye'	(20190710, 2, 25)
*nen.**t**o.ma'*	/nentomaʔ/	'We eat'	(20190708, 3, 34)
*'e.nga.**t**ong*	/ʔeŋatoŋ/	'my/our mouth'	(20190715, 1, 27)
*nen.**t**os.ka'*	/nentoskaʔ/	'our pet animal'	(20190723, 1, 58)
Word-medial onset /k/			
*'as.yay.he.**k**a.ma.ha'*	/ʔasjajhekamahaʔ/	'I run'	(20190717, 2, 65)
*yan.telh.**k**aa.pa'*	/janteɬkaapaʔ/	'December'	(20190715, 1, 57)
*nep.**k**e.sek*	/nepkesek/	'sheep'	(2010716, 1, 57)
*'an.lheng.**k**e'*	/ʔanɬeŋkeʔ/	'You/she goes'	(20190723, 2, 14)
*'as.**k**ok*	/ʔaskok/	'insect'	(20190710, 2, 15)
*poo.**k**o'*	/pookoʔ/	'hot'	(20190710, 3, 50)
Word-medial onset /ʔ/			
*'ak.ve.tay.**'**a*	/ʔakwetajʔa/	'You/she sees'	(20190715, 2, 20)
*yep.ho.pay.**'**a*	/jephopajʔa/	'cloud'	(20190722, 1, 20)
*'ap.na.**'**at*	/ʔapnaʔat/	'your/his face'	(20190722, 2, 50)
*ta.pe.**'**e*	/tapeʔe/	'chicken'	(20190722, 1, 47)
*kel.va.ne.ma.**'**et*	/kelwanemaʔet/	'wolf fish'	(20190710, 2, 11)

Table 2.4: Word-medial onset /p, t, k, ʔ/

The current data do not include combinations of /o/ with coda /p/ or /ʔ/ word medially. /o/ does combine with /k/ and /t/ in word-medial rimes, though less frequently than /a/ and /e/. This is not surprising; /o/ is less common than /a/ and /e/. /ʔ/ also occurs in medial coda position infrequently compared to /p, t, k/. These data also do not include word-medial coda /ʔ/; it may be that /ʔ/ is more restricted or that it is more likely

	Word-medial coda /p/		
*'a**p**.ke.na.o'*	/ʔapkenaoʔ/	'man'	(20190708, 1, 28)
*'em.peng.ka**p**.tek*	/ʔempeŋkaptek/	'my/our elbow'	(20190722, 2, 59)
*ke.la**p**.ha.pa.e'*	/kelaphapaeʔ/	'old'	(20190722, 1, 40)
*'en.ye**p**.he.yok*	/ʔenjephejok/	'our fingers'	(20190722, 2, 6)
*'a**p**.lhep.ko.e'*	/ʔapɬepkoeʔ/	'alone'	(20190715, 2, 70)
	Word-medial coda /t/		
*neng.va**t**.nee.ma'*	/neŋwatneemaʔ/	'It burns'	(20190722, 1, 24)
*'as.ke**t**.ka'*	/ʔasketkaʔ/	'my child'	(20190708, 3, 4)
*'at.ke**t**.so.e'*	/ʔasktsoeʔ/	'small'	(20190710, 3, 37)
*ne.nga.vo**t**.ma'*	/neŋawootmaʔ/	'We weave'	(20190722, 2, 71)
	Word-medial coda /k/		
*'as.ka.me.la**k**.me'*	/ʔaskamelakmeʔ/	'I love (someone)'	(20190719, 1, 7)
*laksa**k**tek*	/laksaktek/	'hominy stew'	(20190722, 2, 15)
*nen.te**k**.may.me.nek*	/nentekmajmenek/	'shoe'	(20190708 2, 76)
*se.le**k**.lek*	/seleklek/	'butterfly'	(2010722, 1, 4)
*'ap.mo**k**.tay.'a*	/ʔapmoktajʔa/	'You/he shoots (something)'	(20190723, 2, 28)

Table 2.5: Word-medial coda /p, t, k, ʔ/

than the other plosives to be elided word-medially without careful pronunciation.

Perhaps more surprising is the difference between /p/ and /t/. /p/ occurs much less frequently as a coda in syllables with /e/ than /t/, and vice versa; /t/ occurs relatively infrequently with /a/ and more often with /e/. As both consonants appear in rimes containing both vowels, this does not appear to be a firm restriction but rather a tendency that prefers /ap/ and /et/ rimes to /at/ and /ep/ ones. The data set also includes a large number of repeated syllables (e.g. *-kek, -mek, -tek, -ma, -yep, -hap, -yet*) which may be separate morphemes whose frequent appearance skews the distribution toward particular combinations of vowels and coda consonants.

2.1.1 Uvular stop /q/

Angaité includes a fifth plosive phoneme, /q/, which is a somewhat exceptional case. First, /q/ in these data appears only adjacent to the low vowel /a/. However, its distribution is not otherwise predictable based on the other stops, and therefore it should be analyzed as a separate phoneme. It appears word-initially and in word-medial syllable onset and coda

position. The current corpus does not include examples of word-final /q/.[1]

Table 2.6 includes most of the examples of /q/ which appear in the current dataset.

	Word-initial /q/		
qaames	/qaames/	'cat'	(20190708, 1, 8)
qala'	/qala?/	'type of duck'	(20190722, 2, 33)
	Word-medial onset /q/		
*'a.ma.**q**a'*	/?amaqa?/	'start, beginning'	(20190708, 2, 1)
*'ang.vang.**q**a.ya.ham*	/?aŋwaŋqajaham/	'congested'	(20190710, 2, 28)
*'a.**q**aa.net*	/?aqaanet/	'two, pair'	(20190710, 3, 13)
*yang.**q**a*	/jaŋqa/	'arrow'	(20190723, 2, 28)
*'ap.**q**aq.hee.ma'*	/?apqaqheema?/	'You/he kills (it)'	(20190720, 1, 9)
	Word-medial coda /q/		
*ye.na**q**.tes.ket*	/jenaqtesket/	'night'	(20190708, 1, 13)
*kel.ya**q**.ha.va.'a*	/kelyaqhawa?a/	'wild pig'	(20190708, 1, 35)
*ka.len.ma**q**.tek*	/kalenmaqtek/	'dove'	(20190708, 1, 63)
*ya**q**.te.pa'*	/jaqtepa?/	'pumpkin, gourd'	(20190708, 2, 9)
*nen.ye.na**q**.te.ma'*	/nenjenaqtema?/	'strong'	(20190708, 2, 94)
*'a**q**.tek*	/?aqtek/	'Your/her eye'	(20190715, 1, 26)
*'ap.qa**q**.hee.ma'*	/?apkaqheema?/	'You/he kills (it)'	(20190723, 2, 26)
*'ak.ya**q**.tan.ma'*	/?akjaqtanma?/	'chainsaw'	(20190723, 1, 31)
*'ap.ya**q**.pas.ka.ma*	/?apjaqpaskama/	'You/he bathe(s)'	(20190723, 2, 12)

Table 2.6: Word-initial and word-medial /q/

This table highlights some notable differences in the distribution of /q/ compared with the other plosives. First, as noted above, /q/ always appears adjacent to /a/, though it can appear as either an onset or coda. Secondly, unlike the other plosives, /q/ is most frequent in syllable codas, but it does not appear word-finally. Finally, /q/ appears very rarely word-initially while other plosives are frequently word-initial.

There are several possible explanations for this distribution. First, I had some difficulty differentiating /q/ and /k/, so it is possible that some things currently analyzed as /k/ are actually /q/. Alternatively, the distinction between /q/ and the /k/ is neutralized

[1]Further support for this analysis of /q/ comes from Unruh and Kalisch (2003), who argue that in all Enlhet-Enenlhet languages except Sanapaná this sound is phonemic. Their classification takes Sanapaná to be Angaité's closest relative, which makes it slightly surprising that Angaité also includes /q/, since the languages are otherwise quite similar. The change from /q/ > /k/ must have occurred independently in Sanapaná.

adjacent to non-low vowels, so we only see /q/ with /a/. Finally, it is possible that /q/ has a wider distribution but is less frequent than the other plosives, and the lack of it in other environments is an accidental gap. Without further testing it is not possible to distinguish between these explanations. Since /q/ clearly contrasts with the other plosives in these data (all of which also appear before and after /a/), I analyze it as phonemic, albeit with a more restricted distribution than the other members of its natural class.

2.1.2 Creaky voice and /ʔ/

The glottal stop is also a somewhat exceptional case. Unruh and Kalisch (1999b) opt not to write glottal stops at word margins, as they are predictable when the word begins or ends with a vowel. Unruh and Kalisch develop an Enlhet orthography, so this is not exactly the same as arguing that glottal stop is epenthetic at word margins. Rather, this decision indicates that Enlhet does not allow word-initial headless syllables nor word-final open syllables. Knowing this, since all other consonants are written in these positions, a word written with an initial or final vowel can be inferred to have an unwritten glottal stop.

Gomes (2013) does, however, argue that word-final V(C) syllables trigger epenthesis in Sanapaná. He states that when headless syllables are underlying at the end of a word, they trigger an epenthetic onset [ʔ] in the final syllable, resulting in two CV.CV syllables. On this analysis, the word-medial [ʔ] in (8) is epenthetic.[2]

(8) a. [paʔa]
'mosquito' (Gomes 2013, p. 128)

b. [jataʔaj]
'goat' (Gomes 2013, p. 128)

c. [aŋaloʔa]
'armadillo' (Gomes 2013, p. 128)

[2]I have reproduced these examples exactly as they appear in Gomes (2013), who does not give underlying forms. However, these phonetic forms are identical to what he seems to propose as the phonemic representation with the exception of the [ʔ], which is epenthetic and therefore not a part of the underlying form.

d. [koʔo]
'I (1SG)' (Gomes 2013, p. 128)

Once again, the argument does not include examples with other morphology that would prevent the CV.V(C) sequence from appearing word-finally in which, presumably, the glottal stop would not appear.

In Angaité, some cases which Gomes would take to be underlying CV.V(C) sequences, as in (8), also have /ʔ/. However, Angaité does not systematically avoid final sequences of CV.V(C). In some cases, rather than a clear glottal stop between the two vowels, what appears is creaky voice on the second vowel. For example, see (9), with creak marked with a tilde under the creaky vowel.

(9) a. *'apkenao'*
[ʔap-kenao̰ʔ]
ʔap-kenaoʔ
2/3M-man
'man' (20190715, 2, 54)

b. *'asvetae'*
[ʔas-witaḛʔ]
ʔas-wetaeʔ
1SG.DIR-see
'I see' (2019715, 2, 33)

c. *'atketsoe'*
[ʔatketsoḛʔ]
ʔatketsoeʔ
'small' (20190710, 3, 37)

This creaky voice may result from the adjacent glottal stop, or may be a prosodic process that occurs regardless of whether or not a glottal stop surfaces as a syllable coda. In support of the latter hypothesis, the same creaky voice that affects the second of two adjacent vowel

segments appears on syllables without glottal stop codas, as in (10).

(10) *heesevaske'*
[he-e̯siwaskeʔ]
he-esewaskeʔ
1SG.INV-love
'It loves me' (20190719, 1, 17)

Here, the elicited corpus currently provides little evidence that the creaky vowel is the result of an underlying glottal stop. Example (10) is multimorphemic; the creaky vowel is the first vowel in what appears to be the verb stem. Future elicitation with different person markers on this stem may clarify whether this creaky voice is a realization of underlying /ʔ/ or a separate phenomenon.

Finally, the word-final /ʔ/ in (9) may be a result of words pronounced in isolation, or it may be part of the same process that results in the creaky vowel in (10). Utterance final creak seems to be a feature of Angaité, with creak or [ʔ] appearing at the ends of utterances and almost all words pronounced in isolation. In some cases, this glottalization disappears when a word is pronounced utterance-medially. When this happens consistently, I do not take the [ʔ] that appears utterance-finally or in isolation to be underlying. Otherwise, I take a conservative approach in this paper and transcribe /ʔ/ when one is pronounced. This means that final /ʔ/ in words which *only* appear in isolation or utterance finally in my data may not turn out to be underlying, but I analyze it as such in lieu of evidence to the contrary. This analysis should be revisited with additional data from elicitation that develops a more robust morphological analysis and compared to running speech in texts. In cases with creaky vowels but no clear glottal stop, I do not transcribe /ʔ/ but rather leave the question open to further investigation of suprasegmental processes in Angaité. The question of glottalization and creak is revisited in §2.2.3 (with nasals) and §2.3.1 (with /h/), and §4.6 (with vowels).

2.1.3 /t/ ∼ /ʔ/ ∼ /k/ alternation

A /t/ sometimes substitutes for /ʔ/ at a morpheme boundary when the following syllable begins with a consonant. Gomes (2013) notes this phenomenon in Sanapaná with

the diminutive suffix *-kok*, where /t/ appears at the end of stems that end in an open syllable. The examples in (11)-(14) with *-ko('o)k* appear in the elicited Angaité dataset, followed by the phonemic form of the stem alone.

(11) a. /kelasmaʔ-koʔok/
[kilasmatkoʔok]
'little fish' (20190722, 2, 27)

b. *kelasma'*
kelasmaʔ
'fish' (20190715, 1, 85)

(12) a. /naataʔ-kok/
[naatatkok]
'little bird' (20190717, 2, 59)

b. *naata'*
naataʔ
'bird' (20190715, 1, 20)

(13) a. /taamaʔ-kok/
[taamatkok]
'little string, thread' (20190722, 2, 69.5)

b. *taama'*
taamaʔ
'string' (20190710, 2, 54)

(14) a. /ʔaaɬʔa-kok/
[ʔaaɬatkok]
'palm heart' (20190719, 1, 34)

b. *'aalha'*
ʔaaɬaʔ
'wax palm' (20190715, 1, 68)

The *-kok* ∼ *-koʔok* alternation marks a plural form on animate stems in Sanapaná (Gomes 2013). In Angaité, however, the *-ko'ok* does not seem to indicate a plural. Plural marking in Enlhet-Enenlhet languages is optional, so it is possible that the plural marking is not salient enough to be consistently translated without a relevant discourse context; the two forms may be more distinguished in naturally occurring discourse.

Because of this stem-final alternation of [ʔ] ∼ [t], in Angaité the [t] cannot unequivocably described as epenthesis as Gomes (2013) proposes for Sanapaná; affixation does not add a segment in Angaité since the stem ends with a consonant in both cases. Rather, [ʔ] substitutes for [t], though the phonetic motivation for this process is unclear.

This [t] between affixes and stems also has a wider distribution in Angaité than what Gomes describes in Sanapaná. It also appears between prefixes and stems that begin with a consonant, as in (15).

(15) /ʔa-semhen/ → [ʔatsemhen]
2/3F-dog
'dog'

In this case, the appearance of [t] has a clearer phonetic motivation. One allomorph of the feminine agreement marker is *'ak-*, so this *'at-* variant likely appears due to place assimilation of the feminine marker's consonant to the initial consonant in the stem. This assimilation occurs sporadically, as the *'ak-* form also appears before alveolars. The prefix-stem alternation may ultimately be related to a different process than the stem-suffix alternation, since in this case [k] and [t] seem to be in free variation while before *-kok* the [ʔ] →[t] change is consistent.

2.1.4 /p/ ∼ /w/ alternation

In one token in my dataset, [p] and [w] alternate. This form is the word for Carandilla Palm, a type of small palm that grows only in the Chaco regions of Paraguay, Argentina, and a small portion of the Mato Grosso do Sul province in Brazil. The forms given by EG and EF as a translation are as follows, with (16a) from EF and (16b) from EG.

(16) a. *va'an*
wa?an
'Carandilla palm'

b. *pa'an*
pa?an
'Carandilla palm'

These pronunciations are consistent, with EF producing *wa'an* at least twice. In one case, EF and EG provided simultaneous translations of the word, where EF used the variant with [w] and EG with [p] (see (20190717, 2, 35)).

Though EG and EF consistently and clearly produced forms with different initial segments, neither form was corrected. Phonemes /p/ and /w/ are contrastive word-initially; the dataset includes the minimal pair in (17) in which both phones contrast before /a/.

(17) a. *vaapa'*
waapa?
'rat' (20190708, 2, 93)

b. *paapa'*
paapa?
'beeswax' (20190710, 3, 35)

In this pair, /p/ and /w/ do not alternate. *Waapa'* cannot mean 'beeswax' and *paapa'* cannot mean 'rat'.

The cause of this [p] ∼ [w] alternation is not clear. Both EG and EF produce /p/

and /w/ in other tokens, so this alternation is not the result of an idiosyncratic phonemic merger in one speaker. Given that EG and EF produce both phonemes in other words, the variation in 'Carandilla palm' may either be due to idiosyncratic variation in that particular term or dialectal variation between the two speakers. The latter hypothesis can be tested with a survey of whether other speakers from different *comunidades* use *pa'an* or *wa'an*.

2.2 Nasals

Angaité has three nasals: /m, n, ŋ/. To determine the phonemic status of these phones, §2.2.1 considers minimal pairs and word-initial and final nasals. Nasal place assimilation and word-medial nasals are considered in §2.2.2, and §2.2.3 discusses word-final nasal glottalization.

2.2.1 Phonemic status of nasals

The dataset includes two (near) minimal pairs that demonstrate these contrasts, one for /m/ ∼ /n/ and one for the /m/ ∼ /ŋ/ contrast.

(18) /m/ ∼ /n/ contrast

a. *meemeh*
meemeh
'mother' (20190722, 1, 33)

b. *neeme'*
neemeʔ
'udder' (20190722, 1, 18)

(19) /m/ ∼ /ŋ/ contrast

a. *lheema'*
ɬeemaʔ
'alone (you, F)' (20190716, 1, 62)

b. *lhenga'*

ɬeŋaʔ

'some species of *Bromelia*' (20190718, 9)

Along with these pairs, the data include contrasting environments for these phonemes at word margins. Both /m/ and /n/ can appear word-initially, though /ŋ/ does not; the absence of word-initial /ŋ/ may be a phonotactic restriction or simply an accidental gap. Table 2.7 shows word-initial /m/ and /n/.

	Word-initial /m/		
maaleng	/maaleŋ/	'fox'	(20190715, 1, 70)
makva'	/makwaʔ/	'peanuts'	(20190722, 2, 16)
metke'	/metkeʔ/	'not, none'	(20190722, 1, 14)
mellhma'	/meɬmaʔ/	'charcoal'	(20190723, 1, 62)
mo'ok	/moʔok/	'other'	(20190708, 3, 35)
momvav'ak	/momwawʔak/	'You don't call me'	(20190722, 2, 26)
	Word-initial /n/		
naata'	/naataʔ/	'bird'	(20190722, 1, 11)
naptekteng	/naptekteŋ/	'frog'	(20190708, 1, 43)
nentoska'	/nentoskaʔ/	'our pet (animal)'	(20190723, 1, 58)
neptaana'	/neptaanaʔ/	'jaguar'	(20190722, 1, 56)

Table 2.7: Word-initial environments for /m, n/

Though /n/ appears word-initially, it does not appear before /o/ in this position. The majority of words that begin with /n/ in my corpus begin with /ne/ as part of the first-person plural marker. As noted in §2.1, /o/ is the least-frequent vowel in the corpus; word-initial /no/ is predicted to appear in a larger sample, given that /n/ can precede /o/ word-medially.

All three phonemic nasals can also appear word-finally, as in Table 2.8. All three vowels appear before word-final /n/ and /ŋ/, but only /e/ and /a/ co-occur consistently with word-final /m/.

		Word-final /m/	
nepyayaam	/nepjajaam/	'Asunción'	(20190708, 2, 27)
talhnaam	/taɬnaam/	'afternoon'	(20190708, 1, 7)
poktem	/poktem/	'timbo'	(20190718)
'aknem	/ʔaknem/	'day, hour'	(20190723, 2, 9)
		Word-final /n/	
lhengvan	/ɬeŋwan/	'ysypo, liana'	(20190723, 1, 37)
'apsan	/ʔapsan/	'toad'	(20190717, 2, 17)
yemen	/jemen/	'water'	(20190708, 2, 97)
leemon	/leemon/	'necklace'	(20190722, 2, 87)
		Word-final /ŋ/	
tomahang	/tomahaŋ/	'rabbit'	(20190723, 1, 41)
melapang	/melapaŋ/	'cactus fruit'	(20190710, 3, 20)
kooneng	/kooneŋ/	'beneath'	(20190710, 3, 8)
metekteng	/metekteŋ/	'duck'	(20190708, 1, 72)
'ektong	/ʔektoŋ/	'my arm'	(20190720, 2, 50)
haapong	/haapoŋ/	'leafcutter ant mound'	(20190718)

Table 2.8: Word-final /m, n, ŋ/

2.2.2 Nasal place assimilation

Nasal place assimilation neutralizes the distinction between nasals when they appear before some sounds, but assimilation does not apply regularly over the entire inventory. The distinction between nasals is always neutralized before /p, t, k/, where they undergo regressive assimilation to the place of articulation of the following stop. This process is most evident in feminine marked forms where the first nasal is always homorganic with the following stop, (20).[3] Relevant segments are **bolded**.

(20) a. *'a**nt**a'ase'*
ʔan-taʔaseʔ
2/3F-beautiful
'beautiful' (20190717, 1, 28)

[3]The underlying form of nasal feminine marker is probably *'ang-*, /ʔaŋ-/ since the velar nasal also appears before vowels. However, the [ŋ] allomorph does not appear entirely consistently. The feminine person marker also has an oral variant, *'ak-*, [ʔak-] and sometimes appears with no following consonant, *'a-*, [ʔa-]. Since /m, n, ŋ/ are all phonemic in Angaité, and the assimilation pattern is not entirely clear, I represent whichever phonemic nasal appears.

b. *'a**mp**ay'a'*
?am-paj?a?
2/3F.DIR-fly
'She/you fly(s)' (20190717, 2, 59)

c. *'a**ngk**amelakme'*
?aŋ-kamelakme?
2/3F.DIR-love
'You/she love(s) (something)' (20190719, 1, 9)

Homorganic nasals also appear word-internally, where no morpheme boundary is clear, (21).

(21) a. *te**mp**eela'*
te**mp**eela?
'type of bird' (20190708, 2, 55)

b. *mo**mp**ehe'*
mompehe?
'type of fish (guaimikue)' (20190715, 1, 22)

c. *ye**nt**apa'*
jentapa?
'firewood' (20190708, 1, 2)

d. *'atsoho' ne**nt**oma'*
'atsoho' nen-toma
sweet 1PL.POS-food
'Our food is sweet' (20190723, 2, 20)

e. *'empe**ngk**aptek*
ʔem-peŋkaptek
1SG.POS-elbow
'my/our elbow' (20190722, 2, 59)

f. *yaasek* *ya**ngk**a'ay* *nenghanma'*
jaasek jaŋkaʔaj neŋ-hanma'
salt season (?) 1PL.POS-food
'We put salt in our food' (20190723, 2, 21)

Though these homorganic nasals may result from assimilation, since they always appear, and /m, n, ŋ/ are contrastive, I treat them as underlying.

Before /j/, a palatal nasal allophone is possible, [ɲ], (22).

(22) a. *'a**ny**engkek*
[ʔaɲjeŋkek]
ʔaN-jeŋkek
2/3F.DIR-drink
'You/she drink(s)' (20190717, 1, 36)

b. *ne**ny**ama'*
[neɲjamaʔ]
neN-jamaʔ
1PL.DIR-drink/eat
'We drink/eat' (20190708, 2, 40)

c. *'a**ny**apong*
[ʔaɲjapoŋ]
ʔaN-japoŋ
2/3F.POS-father
'your/her father' (20190708, 3, 49)

The [ɲ] phone does not appear anywhere except before the palatal glide /j/, so I analyze it as an allophone of the other nasals (capital N above). I phonemically represent the nasal before /j/ as /n/ when it surfaces as [ɲ] (primarily to avoid confusing it with cases in which [ŋ] also surfaces before /j/). In (22), the target nasal is the final nasal of a person marker, but the palatal [ɲ] also appears in monomorphemic tokens, (23).

(23) *pe**ny**et*
[peɲjet]
penjet
'type of fish' (20190723, 1, 39)

Like the examples in (21), I still treat [ɲ] here as underlyingly /n/ which undergoes assimilation to [ɲ]. Because the palatal allophone only appears before /j/, it seems reasonable to posit assimilation here, since the person markers show that assimilation is synchronically active in at least some cases.

However, in a limited number of verbal constructions, place assimilation does not occur, and the nasal before the /j/ surfaces as velar [ŋ], See (24) for examples with *-tingyay'a*, 'to look for'.[4]

(24) a. *'akte**ngy**ay'a* *'aalha'*
[ʔak-tiŋjajʔa ʔaaɬaʔ]
ʔak-teŋjajʔa ʔaaɬaʔ
2/3F.DIR-look.for wax.palm
'You/she look(s) for wax palm' (20190715, 1, 82)

[4]I present here examples with the *-tingyay'a* because it appears multiple times in the corpus, and the [ŋj] sequence appears consistently in each instance; other constructions which have a [ŋj] sequence either only appear once in the corpus or vary in pronunciation between appearances.

b. *nengte**ngg**ay'a* *kelasma'* *'enengko'o'*
[nen-tiŋjaj?a kilasma? ?eneŋko?o?]
neŋ-teŋjaj?a kelasma? ?eneŋko?o?
1PL.DIR-look.for fish 1PL.PN
'We look for fish' (20190715, 1, 85)

One hypothesis for this behavior is that this verb construction may be multi-morphemic, with a morpheme boundary between the nasal and the glide which blocks the spread of assimilation. Since the boundary between the person-markers and the stem does not prevent assimilation, description of categories of verbal morphology is necessary to assess if morpheme boundaries are relevant here. Alternatively, underlying velar nasals may be more resistant to assimilation than /m, n/, a pattern which recurs when other natural classes are considered. This option would help account for forms in the corpus where [ŋ] surfaces where [ɲ] is predicted before /j/. When [ŋ] surfaces in this /_ j environment, I take it to be underlying, rather than /n/.[5]

Just as /j/ irregularly interacts with nasal place assimilation, other inconsistencies in assimilation also surface in these data. Alveolar stops trigger regressive place assimilation for all three nasals, (20), but other alveolar sounds affect /m/ more strongly than /ŋ/. For example, /ŋ/ can surface before /ɬ/, but [mɬ] never does. Compare (25), with a homorganic nasal, with (26), with [ŋ] on the surface.

(25) *ne**nlh**atek*
nen-ɬatek
1PL.DIR-get.up
'We got up already' (20190717, 2, 57)

[5]In other words, because in some words /ŋ/ is maintained before /j/ and in other words [ɲ] appears before /j/, I use /n/ for the underlying phoneme of the nasals which *do* assimilate and /ŋ/ for the underlying representation of those which surface as [ŋ].

(26) *'a**nglh**engke'*
ʔaŋ-ɬeŋkeʔ
2/3F.DIR-go
'You/she go(es) (to Carpincho)' (20190717, 1, 15)

Its interaction with nasal assimilation sets /ɬ/ apart from the other alveolars (/s, t, l/), before which assimilation always occurs. This irregularity provides some basis for considering /ɬ/ to, at least some of the time, pattern as its own natural class.

The /ɬ/ also patterns differently than the other fricatives (/s, h/). Neither /m/ nor /ŋ/ appears before /s/; only /n/ surfaces in this position. However, all three nasals do appear before /h/, (27).

(27) a. *ke**mh**aava'*
kemhaawaʔ
'puma' (20190708, 1, 46)

b. *se**mh**eya'*
semhejaʔ
'type of gourd' (20190708, 2, 25)

c. *ne**ngh**anma'*
neŋ-hanmaʔ
1PL.DIR-boil
'We boil (food)' (20190720, 1, 42)

d. *'e**ngh**aykok*
ʔeŋ-hajkok
1PL.POS-ear
'our ears' (20190715, 1, 29)

e. *'asye**nh**ama'*
ʔas-jenhamaʔ
1SG.DIR-throw
'I throw (to someone)' (20190710, 2, 73)

The fact that /h/ does not trigger nasal assimilation is unsurprising; no nasal is possible at a glottal place of articulation. Together, the three fricatives form a continuum: /s/ which appears only with [n]; /ɬ/ which appears with [n, ŋ] but not [m], and /h/ which never triggers assimilation (appears with [m, n, ŋ]). The four alveolar sounds, as indicated previously, also differ, with /t, s, l/ always triggering assimilation but /ɬ/ only sometimes causing /ŋ/ →[n] assimilation.

Like /h/, /w/ also does not interact with nasal place assimilation. All three nasals appear before /w/, (28).[6]

(28) a. *momvav'ak*
momwawʔak
'You don't call me' (20190720, 2, 26)

b. *'onvetalhka'*
ʔonwetaɬkaʔ
'goodbye' (20190708, 3, 10)

c. *nengvatneema'*
neŋwatneemaʔ
'(It) burns' (20190722, 1,24)

This behavior is more surprising than /h/, since both [m] and [ŋ] are potentially available targets for assimilation for /n/ before /w/. Additionally, /j/ does result in assimilation to [ɲ] in many cases, so the fact that /w/ does not trigger assimilation sets it apart from the

[6]Example (28a) is almost certainly multi-morphemic, but I have only one example of this form in the corpus, and this was the translation provided for it. This may also only be the negator (with some T/A/M marking?), and not the verb construction itself, but the datum isn't clear.

other glide.

These three exceptional cases—/w, h, ɬ/—show an asymmetrical system of nasal assimilation. Nasals assimilate to the place of articulation of the following consonant, except for before /h, w/, where they contrast. Furthermore, the contrast between /n/ and /ŋ/ is maintained before /ɬ/, but there is no contrast between /m/ and /n, ŋ/ in this position. Behaving differently from the other alveolars, /ɬ/ does not cause /ŋ/ assimilation. It also differs from the other fricatives by causing assimilation for only one nasal phoneme, unlike /s/ which always triggers assimilation and /h/ which never does. And, finally, /j/ sometimes, but not always results in the appearance of [ɲ]; [ŋ] is also possible before [j].

In monomorphemic words, or multimorphemic words which are always attested in these data in the same constructions, it is impossible to determine whether place assimilation is an active, synchronic process or an historical one. The person markers provide evidence that at least some place assimilation is synchronically active, since these markers have a variety of environmentally-conditioned allomorphs. However, most word-medial nasals always surface the same way, so the process could be either active or static based on these data. Because of this uncertainty, phonemic representation of nasals in this paper corresponds to their phonetic realization unless the phonetic realization is [ɲ], in which case it is phonemically represented as /n/.

All three phonemic nasals appear word-medially, either as onsets or codas, as in Table 2.9 for onset position and Table 2.10 for coda position. The phoneme /ŋ/ only appears as an onset before /a/ in these data. The explanation may be the same here as for word-initial /no/ combinations, discussed above. However, this dataset is fairly small, and onset /ŋe/ and /ŋo/ clusters may appear in a larger sample. Presently, we can say that /ŋ/ has a more restricted distribution than /m/ and /n/, appearing neither word-initially nor as a word-medial onset before /e, o/.

	Word-medial onset /m/		
*'ak.tay.ka.**m**a.ha*	/ʔaktajkamaha/	'You/she work(s)'	(20190710, 2, 67)
*'a.pan.**m**ah*	/ʔapanmah/	'You/he sow(s)'	(20190719, 1, 30)
*ka.**m**e.lang.ko.ho'*	/kamelaŋkohoʔ/	'slowly'	(20190723, 2, 7)
*talh.nak.**m**e'*	/talnakmeʔ/	'afternoon'	(20190717, 2, 54)
*mee.**m**ong*	/meemoŋ/	'palo santo'	(20190708, 2, 2)
*lee.**m**on*	/leemon/	'necklace'	(20190722, 2, 87)
	Word-medial onset /n/		
*nen.ye.**n**aq.te.ma'*	/nenjenaqtemaʔ/	'strong'	(20190708, 2, 94)
*talh.**n**ak.me'*	/talnakmeʔ/	'afternoon	(20190717, 2, 54)
*'a.me.**n**ek*	/ʔamenek/	'her calf (body part)'	(20190716, 1, 35)
*so.mo.**n**e.ye'*	/somonejeʔ/	'watermelon'	(20190708, 2, 8)
*'ak.yelh.**n**o.yay.ka.ha*	/ʔakjelnojajkaha/	'year/December'	(20190710, 3, 17)
	Word-medial onset /ŋ/		
*ne.**ng**a.voot.ma'*	/neŋawootmaʔ/	'We weave'	(20190722, 2, 71)
*'ak.lhe.**ng**a.ma'*	/ʔakleŋamaʔ/	'You/she walk(s)'	(20190710, 2, 84)
*lhe.**ng**a'*	/leŋaʔ/	'some species of *Bromelia*'	(20190718)

Table 2.9: Word-medial onset /m, n, ŋ/

	Word-medial coda /m/		
*'e.ma.lhe**m**.pe.nek*	/ʔemalempenek/	'my cheek'	(20190710. 2, 42)
*'a**m**.pe.nek 'apyesvasema'*	/ʔampenek ʔapjeswasemaʔ/	'kururu pytã'	(20190717, 2, 18)
*lhen.ta**m**.pe'*	/lentampeʔ/	'widower'	(20190715, 1, 19)
*se**m**.hen*	/semhen/	'dog'	(20190708, 3, 7)
*ke**m**.haa.va'*	/kemhaawaʔ/	'puma'	(20190708, 1, 46)
*mo**m**.pe.he'*	/mompeheʔ/	'guaimikue' (type of fish)	(20190715, 1, 22)
*te.mo**m**.'a'*	/temomʔaʔ/	'karaguatá'	(2019717, 2, 32)
*mo**m**.vav.'ak*	/momwawʔak/	'You don't call me'	(20190720, 2, 26)
	Word-medial coda /n/		
*'as.ka**n**.mah*	/ʔaskanmah/	'I sow'	(20190708, 2, 28)
*'a**n**.ta.va'*	/ʔantawaʔ/	'poroto del monte'	(20190708, 2, 6)
*'a**n**.mo.wa.ma'*	/ʔanmowamaʔ/	'rifle'	(20190723, 2, 30)
*ye**n**.taa.pa'*	/jentaapaʔ/	'firewood'	(20190719, 1, 36)
*'as.ye**n**.ha.ma*	/ʔasjenhama/	'I throw (to someone)'	(20190710, 2, 73)
	Word-medial coda /ŋ/		
*'a**ng**.ken*	/ʔaŋken/	'your/her mother'	(20190722, 1, 34)
*'a**ng**.lhe**ng**.ke'*	/ʔaŋleŋkeʔ/	'You/she go(es)'	(20190717, 1, 15)
*nen.te**ng**.yay.'a*	/nenteŋjajʔa/	'We look for (something)'	(20190715, 1, 82)
*pas.ko**ng**.kong*	/paskoŋkoŋ/	'ant'	(20190708, 2, 90)
*'a.so**ng**.ko.e'*	/ʔasoŋkoeʔ/	'ugly (?)'	(20190723, 2, 16)

Table 2.10: Word-medial coda /m, n, ŋ/

2.2.3 Creaky voice and nasals

Some sources on Enlhet-Enenlhet languages indicate that word-final nasals (and perhaps other sonorants) are glottalized. Key and Comrie (2015), based on Unruh and Kalisch (1997) (listed as Sanapaná) includes glottalization on some word-final nasals and glides, and a glottal stop coda in some words; some other words end with plain nasals or glides, or open syllables. In this case, the word-final glottalization is not predictable and is included in a phonemic transcription provided by the list. In contrast to their previous work, Unruh and Kalisch (1999b) state that a glottal stop always appears after word-final nasals and vowels in Enlhet. They choose not to represent this segment in the orthography because its appearance is predictable.

Unruh and Kalisch's evolving analysis of Enlhet presents two potential options to examine in Angaité. The earlier system used in the Enlhet dictionary and Key and Comrie (2015) indicates that sometimes word-final nasals and glides are glottalized, while sometimes sonorants, including vowels, can appear word-finally with no glottalization. Kalisch and Unruh subsequently altered the Enlhet orthography to reflect a revised analysis that takes glottalization and glottal stops at word-margins to be entirely predictable. Despite these differing stances, these two analyses do converge on one point: glottalized sonorants appear only word-finally and not word-medially.

As discussed in §2.1.2, §2.3.1 and §4.6, creaky voice appears in Angaité both word-finally and word-medially. Unruh and Kalisch's analysis of Enlhet does not consider word-medial creak; one issue in the analysis of Angaité, then, is determining whether word-medial creak is governed by the same process(es) as word-final glottalization. Creak in the Angaité data may be due to a variety of sources: suprasegmental processes at the word or utterance level, where creak seems to always appear; underlying /ʔ, h/ realized as phonation and pitch changes; or factors specific to the production of words in isolation. Impressionistically, some words that end with nasals are creakier than others. This creak is reflected in measures of local jitter for the final rime (vowel + nasal), calculated by Praat as the difference between consecutive periods (glottal pulses) divided by the average period. Compare Figure 2.1, the spectrogram for the final rime of *'aknem* 'day' and Figure 2.2, for *'aptomahang* 'rabbit'.

Local jitter for the rime in *'aknem* is 2.393% compared to 1.096% in *'aptomahang*. Visually, this higher amount of creak in *'aknem* can be observed in the more irregular spacing of the blue bars marking glottal pulses on the waveform.

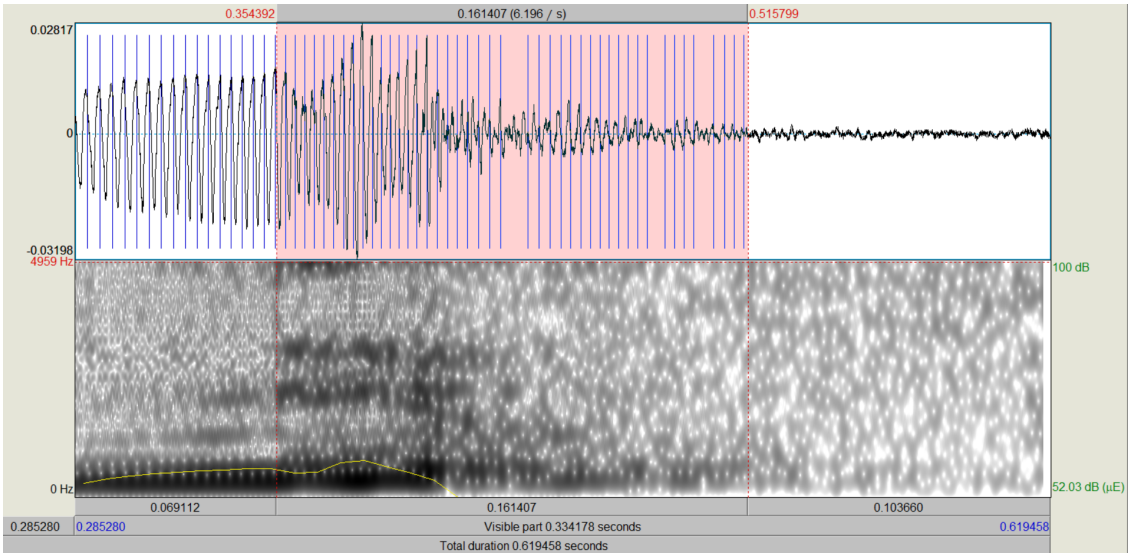

Figure 2.1: Spectrogram for *em* in *'aknem*, produced by EF (20190708, 1, 71)

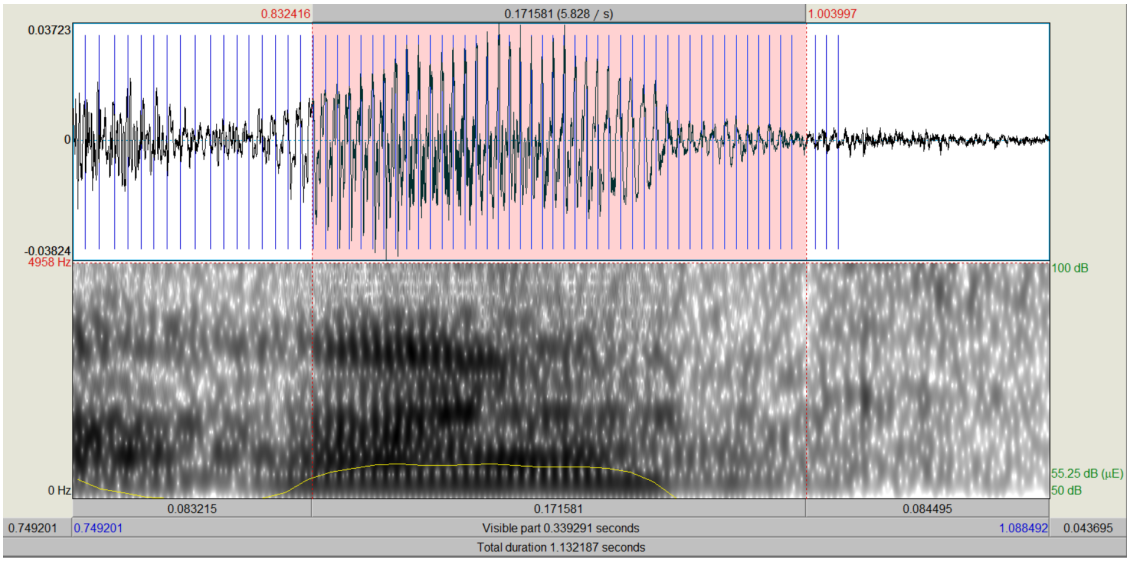

Figure 2.2: Spectrogram for *ang* for *'aptomahang*, produced by EF (20190710, 2, 8)

Further complicating this picture is that van Gysel (2017) does not mention glottalization of sonorants in his discussion of Sanapaná, though he does discuss glottalization of /ɬ/ which does not appear in other Enlhet-Enenlhet languages. These various treatments of glottalization indicate both that the phenomenon varies, perhaps substantially, within the family, and that it interacts with word or phrase-level prosody, which is beyond the scope of this paper. At present, the Angaité data provide clear evidence for the three phonemic nasals /m, n, ŋ/ and suggest some interaction between these segments and phonation which may be (or not) contrastive.

2.3 Fricatives

The three phonemic fricatives in Angaité are /s/, /ɬ/, and /h/. The data do not contain minimal pairs for these contrasts. There are, however, (near) minimal pairs showing that both alveolar fricatives contrast with /t/, as in Table 2.11.

*'akve**t**ay'a'* /ʔakwetajʔaʔ/	'You/he see(s)'	(20190717, 1, 11)	*'akve**s**ay'a* /ʔakwesajʔaʔ/	'You/she are/is named'	(20190710, 3, 32)
*taa**t**a'* /taataʔ/	'father'	(20190710, 2, 17)	*taa**lh**a'* /taaɬaʔ/	'fire'	(20190708, 1, 1)

Table 2.11: (Near) minimal pairs showing alveolar contrasts

Though there are not minimal pairs in this dataset that demonstrate the contrast between the fricatives, as in other natural classes, the three phonemes appear in contrasting environments (word-initial, word-final, word-medial onset and coda). All three sounds appear word-initially, in Table 2.12.

		Word-initial /s/	
saangah	/saaŋah/	'lagoon'	(20190722, 1, 6)
kelyapetsanake'	/keljapetsanakeʔ/	'vulture'	(20190708, 1, 68)
sesenhe'	/sosenheʔ/	'net bag'	(20190710, 3, 48)
sepo'	/sepoʔ/	'manioc'	(20190708, 2, 12)
somoneye'	/somonejeʔ/	'watermelon'	(20190720, 1, 37)
sosekha'	/sosekhaʔ/	'tomorrow'	(20190717, 1, 7)
		Word-initial /ɬ/	
lhalhanha'	/ɬaɬanhaʔ/	'together'	(20190717, 1, 7)
lhamakha'	/ɬamakhaʔ/	? something to do with reciprocals?	(20190723, 1, 53)
lheya'	/ɬejaʔ/	'2/3F	(20190723, 1, 53)
lhepop	/ɬepop/	'dust'	(20190722, 1, 50)
lheema'	/ɬeemaʔ/	'alone'	(20190716, 1, 62)
		Word-initial /h/	
hasnengakok	/hasneŋakok/	'(It) hears (me)'	(20190719, 1, 20)
haapong	/haapoŋ/	'small white fish-eating bird'	(20190722, 2, 26)
heva'	/hewaʔ/	'spear'	(20190710, 3, 41)
heta'aye'	/hetaʔajeʔ/	'You/she see(s) (me)'	(20190715, 2, 21)

Table 2.12: Word-initial /s, ɬ, h/

Fricatives can also appear word-medially. /s, ɬ/ can appear either as onsets or codas, but /h/ appears only as an onset word-medially. Since /h/ patterns like /s, ɬ/ elsewhere, it may simply be less frequent in this position and could appear there in a larger sample. Alternatively, since /h/ is a weak consonant compared to /s, ɬ/ it may just not appear in word-medial codas. If the latter is true, one could likely refer to an historical explanation for this gap in the distribution, since deletion of /h/ is a common diachronic change. Table 2.13 shows all three fricatives as word-medial onsets and Table 2.14 shows them as medial codas.

Word-medial onset /s/			
*'a.**s**an.mah*	/ʔasanmah/	'I sow'	(20190708, 3, 33)
*hat.**s**ap*	/hatsap/	'toucan'	(20190718)
*'a.lat.**s**e.he'*	/ʔalatsehe?/	'corn'	(20190708, 2, 16)
*'at.**s**e.ta'*	/ʔatseta?/	'passion flower'	(20190717, 2, 33)
*'at.**s**o.**s**e.ko.e'*	/ʔatsosekoe?/	'early morning'	(20190720, 1, 54)
*'ak.ma.**s**o.ma*	/ʔakmasoma/	'You/she are/is ugly'	(20190710, 3, 81)
*nek.**s**o.ma'*	/neksoma?/	'grasshopper'	(20190710, 2, 19)
Word-medial onset /ɬ/			
*'a.**lh**ang.kok*	/aɬaŋkok/	'your/her house'	(20190722, 1, 44)
*'ap.kel.yep.**lh**ay.'a*	/ʔapkeljepɬajʔa/	'You/he cut (something)	(20190710, 3, 42)
*'ak.**lh**e.nga.ma'*	/ʔakɬeŋama?/	'You/she walk(s)'	(20190710, 2, 84)
*'an.ya.**lh**o'*	/ʔanjaɬo?/	'your/her sister'	(20190719, 1, 9)
Word-medial onset /h/			
*'a.**h**an.kok*	/ʔahankok/	'your/her house'	(20190716, 1, 26)
*nen.ta.**h**ay.ka.ma.**h**a*	/nentahajkamaha/	'We work'	(20190717, 1, 6)
*'ak.yey.**h**ee.ma'*	/ʔakjejheema?/	'You/she dance(es)'	(20190710, 2, 76)
*'ap.yem.pe.**h**ek*	/ʔapjempehek/	'your/his skin'	(20190720, 2, 46)
*ko.ya.ye.**h**o.**h**o'*	/kojajehoho?/	'fast'	(20190723, 2, 6)
*'at.so.**h**o'*	/ʔatsoho?/	'sweet'	(20190708, 2, 52)

Table 2.13: Word-medial onset /s, ɬ, h/

Word-medial coda /s/			
*'a.nayk.ma**s**.ka.ma'*	/ʔanajkmaskama?/	'You/she sing(s)'	(20190716, 1, 14)
*'ak.me.a**s**.ma'*	/ʔakmeasma?/	'You/she speak(s)'	(20190722, 1, 25)
*'e.ya**s**.pok*	/ʔejaspok/	'my neck'	(20190710, 2, 34)
*ye.nak.te**s**.ke'*	/jenakteske?/	'night'	(20190720, 1, 45)
*'as.to**s**.ka'*	/ʔastoska?/	'my (pet) animal'	(20190708, 3, 45)
*mo.**s**ak.ha'*	/mosakha?/	'I/you/we don't bring'	(201970820, 2, 30)
Word-medial coda /ɬ/			
*ta**lh**.nak.me'*	/taɬnakme?/	'afternoon'	(20190722, 2, 58)
*'a.ka**lh**.kok*	/ʔakaɬkok/	'your/her tongue'	(20190722, 2, 58)
*pe**lh**.mok*	/peɬmok/	'grease, fat'	(20190710, 2, 53)
*ke**lh**.vo.ye'*	/keɬwoje?/	'soon'	(20190723, 2, 11)
*yan.te**lh**.kaa.pa'*	/janteɬkaapa?/	'year'	(2019010, 3, 16)
*'o**lh**.te.nak.sek*	/ʔoɬtenaksek/	'I will tell (someone)'	(20190716, 1, 43)

Table 2.14: Word-medial coda /s, ɬ/

All three fricatives also appear word-finally, though this position is infrequent compared to the plosives and nasals; see Table 2.15. In these data the fricatives are less frequent

		Word-final /s/	
qaames	/qaames/	'cat'	(20190722, 1, 54)
teeves	/teewes/	'Jungleplum'	(20190715, 1, 3)
		Word-final /ɬ/	
kolayelh	/kolajeɬ/	'easy'	(20190710, 3, 72)
pelhpa'alh	/peɬpaʔaɬ/	'bow'	(20190723, 2, 27)
		Word-final /h/	
lalyah	/laljah/	'dew'	(20190710, 3, 3)
'askanmah	/ʔaskanmah/	'I plant'	(20190720, 1, 33)
yalvah	/jalwah/	'three-banded armadillo'	(20190708, 1, 61)
saangah	/saaŋah/	'swamp'	(20190710, 1, 7)

Table 2.15: Word-final /s, ɬ, h/

than the plosives and nasals, and their distributions are also more restricted. /ɬ/ appears more than /s/ in word-medial coda position, and /h/ is more restricted, as it never appears as a word-medial coda. Currently, /s/ is only attested word-finally following /e/, and /h/ only after /a/. /ɬ/ appears word finally with both /e, a/, but not /o/. Given that the fricatives are less frequent than the stops and nasals, I expect that these apparent restrictions on word-final fricative-vowel combinations are an accidental gap. All three fricatives also have slightly different behavior with respect to nasal assimilation, discussed in §2.2.2, with /h/ never causing place assimilation, /s/ always causing it, and /ɬ/ apparently triggering assimilation for /m/ but not /ŋ/.

2.3.1 Creaky voice and /h/

Similar to /ʔ/ discussed in §2.1.2, /h/ also alternates with creaky voice on adjacent vowels, particularly word-medially. In careful pronunciation, /h/ is clearly pronounced, but in faster or more connected speech, it is often realized as creak on the following vowel, as in (29a), compared to (29b), pronounced one after the other.

(29) a. *pomahap*

[pomaa̰p]
/pomahap/
'pig' (20190708, 3, 42, EF)

b. *pomahap*

[pomahap]
/pomahap/
'pig' (20190708, 3, 42, EF)

Sometimes this realization is also accompanied by syllable deletion, where the creaky vowel is the only thing that remains from the deleted syllable. This is the case in (30a), where the syllable with /h/ is deleted, and the final vowel is realized with creak. Compare to (30b) where the second syllable is maintained in careful speech.

(30) a. *peheya'*

[pija?]
/peheja?/
'sweet potato' (20190708, 2, 11, EF)

b. *'askanmah peheya'*

/?as-kanmah peheja?/
1SG.DIR-sow
'I sow sweet potato' (2019708, 2, 31, EG)

EG seems to be more prone to this syllable deletion and realization of /h/ as creaky voice, which suggests that it may partially be an age or dialect related feature. However, both EF and EG do produce these alternations, indicating that the variation is not entirely idiosyncratic. In this way, /h/ patterns similarly to /?/; both glottal consonants are realized as creaky voice in rapid pronunciation. However, determining the underlying representation here is slightly more straightforward than with /?/, since utterance-final glottalization cannot be mistaken for an underlying /h/ in careful speech.

2.4 Laterals

Like some other Chaco languages (Mataco-Mataguayo family and Vilela (Lule-Vilela), see Campbell, 2012; 2013), Angaité's inventory includes a class of laterals which contrast based on mode of articulation. Some authors have analyzed this feature as propagated by contact between Chaco languages, though in the case of Enlhet-Enenlhet languages shoring up this argument requires more descriptive and historical work than is currently available. The voiceless lateral fricative was discussed along with other fricatives in §2.3. This section deals with /l/, which patterns differently enough from the other approximants to warrant a separate discussion.

Both laterals occur relatively infrequently compared to other consonants, especially the plosives and nasals, though /ɬ/ is more common than /l/. It appears at the beginning of words, as in Table 2.16. In my data, /l/ appears more frequently before /a/, while /ɬ/

latsehe'	/latsehe?/	'corn'	(20190720, 1, 31)
lalyah	/laljah/	'dew'	(20190710, 3, 3)
lahak	/lahak/	'hook'	(20190710, 3, 42)
leemon	/leemon/	'necklace'	(20190722, 2, 87)

Table 2.16: Word-initial /l/

more frequently appears before /e/. This corpus is too small to determine whether this is a broad pattern in the language or an incidental sampling bias. Furthermore, /ɬ/ does appear before /a/ just as /l/ does appear before /e/, so the distributions are not complementary.

Like /ɬ/, /l/ also appears word-medially either as a syllable onset or a coda, as in Table 2.17. Unlike /ɬ/, however, /l/ never appears word-finally. Borrowing accommodations indicate that this absence is not an accident of the sample. For example, the personal name 'Marcial' is pronounced [marsja] or [masja], with the final /l/, which is pronounced in Spanish, eliminated.

Other sonorants (nasals and glides) frequently do appear word-finally, which makes it surprising that /l/ never occurs in this position. One option is that word-final devoicing causes a neutralization of /l/ and /ɬ/ in that position. This analysis predicts that some

	Word-medial onset /l/		
*'as.ka.me.**l**ak.me'*	/ʔaskamelakmeʔ/	'I love (something)'	(20190719, 1, 7)
*tem.pee.**l**a'*	/tempeelaʔ/	'type of water bird'	(20190722, 1, 7)
*maa.**l**eng*	/maaleŋ/	'fox'	(20190715, 1, 70)
*se.**l**ek.**l**ek*	/seleklek/	'butterfly'	(20190722, 1, 4)
*'ak.**l**oo.ma*	/ʔaklooma/	'You/she is/are bad'	(20190710, 3, 68)
*'as.**l**oo.ma*	/ʔaslooma/	'I am bad'	(20190719, 1, 45)
	Word-medial coda /l/		
*'a**l**.yas.kes.ka.ma'*	/ʔaljaskeskamaʔ/	'You/she enter(s)'	(20190716, 1, 52)
*ya**l**.vah*	/jalwah/	'three-banded armadillo'	(20190715, 1, 74)
*ke**l**.va.na*	/kelwana/	'woman'	(20190708, 1, 29)
*'ang.ke**l**.yak*	/ʔaŋkeljak/	'You/she boil(s) (something)'	(2019010, 2, 51)
*mo**l**.yas.kaa.lhak*	/moljaskaaɬak/	'You/she do(es) not wash'	(20190720, 2, 8)

Table 2.17: Word-medial /l/

forms that end in /ɬ/ should show /l/ at the end of the stem when an additional suffix is added. For example, we might predict (31).

(31)** [peɬpaʔaɬ] + /-kok/
bow + DIM
[peɬpaʔalkok]
'little bow'

But, it does not explain why the /l/ is deleted in borrowings rather than changed to /ɬ/. Alternatively, since /l/ is relatively infrequent compared to the other sonorants, it may just have a slightly more restricted distribution. Future elicitation based on a more comprehensive morphological description should be able to work this out.

2.5 Glides

Angaité also has two phonemic glides, /w/ and /j/. The evidence for phonemic glides comes from glides in onset position, where they contrast with the other consonants. The elicited data contains a near-minimal quadruplet for this contrast: *yaapa'* 'money' ∼ *vaapa'* 'rat' ∼ *paapa'* 'beeswax' ∼ *taata'* 'father'. For examples of the glides in contrasting word-initial environments, see Table 2.18.

		Word-initial /j/	
yapa'	/jaapaʔ/	'money'	(20190708, 2, 89)
yamapataymenek	/jamapatajmenek/	'sugar'	(20190708, 2, 50)
yaamet	/jaamet/	'tree'	(20190708, 2, 96)
yetalheng	/jetaɬeŋ/	'horse'	(20190716, 1, 25)
yetayavhan	/jetajawhan/	'honey'	(20190723, 2, 22)
yemen	/jemen/	'water'	(20190708, 2, 97)
		Word-initial /w/	
vaapa'	/waapaʔ/	'rat'	(20190722, 1, 54)
valayo'	/walajoʔ/	'foreigner'	(20190708, 3, 61)
vaavo'	/waawoʔ/	'maned wolf'	(20190717, 2, 3)
venak	/wenak/	'right now'	(20190722, 2, 38)
velhaasek	/weɬaasek/	'type of many legged bug'	(20190720, 2, 38)
vona'	/wonaʔ/	'net bag' (borrowed from Guaraní)	(20190710, 2, 70)

Table 2.18: Word-initial /j, w/

This dataset does not include any examples of word-initial glides preceding /o/ in native Angaité words. The only example of a glide before /o/ appears in *vona'*, 'net bag', which is borrowed from Guaraní (the Angaité word for this item is *sosenhe'*[7]). As noted in the discussion of other consonants, /o/ is generally less common than /a/ and /e/, especially in initial syllables because /o/ does not appear in person markers.

Glides also contrast with other consonants in word-medial onset position, as in Table 2.19. I do not analyze sequences of a vowel followed by a glide as diphthongs (see Chapter 3), but to avoid any ambiguity, the examples here are ones in which the glide is unequivocally an onset (i.e. following another consonant) and cannot be interpreted as part of the preceding nucleus. The /j/ is unique as the only palatal sound in the inventory, and as such plays a special role with respect to its relationship to nasals, where it conditions a nasal allophone, [ɲ], the only non-phonemic nasal allophone. As described in §2.2.2, all other cases of nasal place assimilation result in neutralization rather than an allophone that never appears as its own phoneme.

Though glides appear to pattern like other consonants in onset position, there is a history of theoretical debate in phonological literature over whether glides and high vowels

[7]This word is one which has a vowel alternation. It is sometimes pronounced [sɛsɛnhɛʔ]; the pronunciation with /o/ is more frequent, so I take it to be underlying.

	Word-medial onset /j/		
*'ap.**y**a.ma'*	/ʔapjamaʔ/	'You/he drink(s)'	(20190722, 2, 24)
*'as.teng.**y**a.'ay*	/ʔasteŋjaʔaj/	'I search (for something)'	(20190719, 1, 36)
*'ap.**y**em.pe.hek*	/ʔapjempehek/	'your/his skin'	(20190720, 2, 46)
*'ap.**y**eng.kek*	/ʔapjeŋkek/	'You/he drinks (mate)'	(20190717, 1, 33)
	Word-medial onset /w/		
*'ang.kel.**v**a.na*	/ʔaŋkelwana/	'woman'	(20190722, 1, 30)
*'al.**v**a.ta'*	/ʔalwataʔ/	'saltwater pond, river'	(20190708, 2, 38)
*'as.**v**e.ta.ye'*	/ʔaswetajeʔ/	'I see'	(20190708, 3, 38)
*'on.**v**e.talh.ka'*	/ʔonwetałkaʔ/	'goodbye/see you later'	(20190708, 3, 10)
*kelh.**v**o.ye'*	/kełwojeʔ/	'soon/fast'	(20190723, 2, 11)

Table 2.19: Word-medial onset /j, w/

should be considered separate phonemes or positional variants of one another with syllable structure or prosody determining which surfaces. Therefore, in order to argue that glides are phonemic they should not only pattern like other consonants but also their distribution should not be predictable based on the distribution of vowels, particularly high vowels. This question is taken up here, where I ultimately conclude that glides and high vowels are separate phonemes.

Levi (2011) provides a typology of vocoid systems, presented in Table 2.20, with eight logically possible options, not all of which are attested in natural language. In Types I through IV, only one phoneme is underlying, either /G/ or /V/, with either one or two allophones [V], [G], or both. Types V through VIII all have two underlying phonemes.

One-phoneme systems		Two-phoneme systems		
Type I	/V/ ∼ [V]	Type V	/V/ ∼ [V]	/G/ ∼ [G]
Type II	/G/ ∼ [G]	Type VI	/V/ ∼ [V, G]	/G/ ∼ [G]
Type III	/V/ ∼ [V, G]	Type VII	/V/ ∼ [V]	/G/ ∼ [G, V]
Type IV	/G/ ∼ [G, V]	Type VIII	/V/ ∼ [V, G]	/G/ ∼ [G, V]

Table 2.20: Typology of vocoid systems, from Levi (2011)

Per Levi, Types IV and VII are unattested. As an explanation, she proposes an implicational universal that glides may not alternate with high vowels if phonemic vowels do not alternate with glides, which accounts for the absence of both Types IV and VII.

The system that I propose for Angaité is Type II, where glides are underlying and

appear only as glides, and there is no alternation with high vowels. Though some high vowel phones do appear in Angaité, they are not in complementary distribution with glides but rather are allophones of mid vowels /e, o/. At first glance, glides and high vowels do appear to have separate environments: glides can appear word-initially and vowels never do, and high vowels appear as nuclei, which glides never do. However, VV sequences (hiatus) appear with [i] in combination with /a/, (32).[8]

(32) a. *nent**ea**nma*
[nentianma]
nenteanmaʔ
'We sleep' (20190710, 3, 78)

b. *'akmam**ea**y'a'*
[ʔakmamiajʔaʔ]
ʔakmameajʔaʔ
'It rains' (20190722, 1, 22)

We also get [ja] sequences in these same environments. See (33) where [j] appears in two environments: /C$_{[stop]}$_a, (33a), and /N_a, (33b).

(33) a. *'ap**y**ayaam*
[ʔapjajaam]
'south wind, winter' (20190723, 2, 17)

b. *'en**y**alho'*
[ʔeɲjałoʔ]
ʔen-jałoʔ
1SG.POS-older.sibling
'my older sibling' (20190722, 1, 38)

[8]Vowel hiatus is its own acoustic question which merits some closer analysis at a later date. The cases that I call hiatus here are those in which (a) I clearly heard two syllables and, more importantly, (b) transitions between the two vowels on a spectrogram are fairly direct, with no sign of a glide indicated in either formant transitions or decreased intensity.

One might argue that the VV sequences are only allowed when realizing one of the mid Vs as a glide would create a disallowed consonant cluster. This rule works for (32a), where realizing [i] as [j] would create a tautosyllabic consonant cluster. However, in (32b), [j] could appear instead of [i] without causing a disallowed cluster. Since GV sequences appear in similar environments, predicting this vowel ∼ glide alternation is impossible, indicating that glides and high vowels belong to separate phonemes.

If this examination of vowel ∼ glide alternations is expanded to include vowel hiatus with mid vowels, (34) and (35) are also comparable.

(34) *m**ae**k*
[maek]
'hunger' (20190722, 2, 1)

(35) *'a**nayk**maskama'*
[ʔanajkmaskamaʔ]
'You/she sings' (20190719, 1, 42)

In (35), as discussed further in Chapter 3, the sequence [najk] creates either an exceptional coda consonant cluster [jk] or a diphthong. If this syllable underlyingly contained a VV sequence [ae] the appearance of [j] instead of [e] would be highly unexpected, since a surface [e] with vowel hiatus would not produce an irregular cluster. This option would rather create a sequence of [NV.Vk] identical to what appears in (34). Since both options are available in Angaité, the only explanation is that (34) contains two underlying vowel phonemes and (35) contains a vowel-consonant sequence.[9]

[9]Furthermore, (34) functions similarly to other bisyllabic words with two full vowels in that the second vowel has a higher pitch. I compared both vowels in *maek* with two other bisyllabic words with similar shapes (*yaayet* 'gourd', *yaamet* 'tree') and found that the second vowel was higher pitched in all cases by between 20 and 80 Hz. I also measured vowel duration for these three words and three other words which do not appear to have long vowels (*akpet* 'cactus fruit', *mepop* 'bat' and *tahan* 'fish hook') and found that in all cases the first vowel was longer than the second, though the difference in *maek* is about 8ms, which is negligible. Over two repetitions each of *'anaykmaskama'* and *maek*, *maek* is also 99.23ms longer on average than the *najk* syllable. This duration and pitch information should be followed up with more controlled recordings before being considered definitive but as a first pass indicates that *maek* is two syllables with vowel hiatus, while *nayk* in *'anaykmaskama'* is one.

Further support for treatment of glides as phonemes separate from the vowels comes from nasal place assimilation. The [j] often triggers nasal assimilation to [ɲ]. If [j] is a positional variant of [i], we might expect [i] to also cause nasal assimilation to [ɲ]. However, this does not occur, (36).

(36) a. *'aanek*
[ʔaanik]
ʔaanek
'seed' (20190710, 3, 18)

b. *hasnengakok*
[hasniŋakok]
hasneŋakok
'It hears (me)' (20190719, 1, 20)

c. *'aqaanet*
[ʔaqaanit]
ʔaqaanet
'two' (20190716, 1 , 51)

These pieces of evidence taken together—nasal place assimilation with [j] but not [i], unpredictable vowel hiatus, and distribution of high vowels compared to mid vowels—suggest that glides before vowels are not positional variants of high vowels but rather separate phonemes that pattern like all other consonants in Angaité.

Chapter 3

Syllable Structure

Though glides are phonemic consonants in Angaité, their status in VG sequences is not so clear. One possibility is that glides in VG sequences are the offglide of falling diphthongs. Alternatively, glides in VG sequences may be coda consonants, forming VC sequences. Both analyses result in glides participating in heavy syllables in Angaité, either CVV syllables or CVC ones. Both options also create some irregularities in allowable syllable structures and consonant distributions within syllables.

Typological literature on syllable structure provides some context toward untangling this problem in Angaité. The literature reviewed here suggests a variety of avenues for future research that would clarify some of the issues arising here. In this way, the typological literature informs my future fieldwork trajectory perhaps more than it directs the conclusions of this particular project.

3.1 Evidence for Syllables

The broadest issue related to syllables is whether the syllable is a universally relevant (or present) unit. Hyman (2010) presents five possible arenas in which one might find evidence for the syllable and argues that evidence in at least one of these is necessary in a given language: (1) distributional constraints governed by syllable structure, (2) phonological or (3) morphological rules conditioned by syllable structure, (4) prosodic processes or word-stress that target the syllable, and (5) prosodic grouping of syllables into higher-order constituents like feet.

Without a more thorough analysis of Angaité morphology and syntax, evidence of morphological rules conditioned by syllable structure is not apparent. However, consonant distribution in Angaité does provide some evidence for syllable structure. Regardless of how

VG sequences are analyzed, Angaité allows no more than two consonants per syllable (with a very small set of possible exceptions, see §3.5.1). Furthermore, within a syllable consonants may not be adjacent, though they may be juxtaposed across syllable boundaries.

Prosody also presents some evidence for the relevance of the syllable. Stress is primarily word-final, with the exception of some affixes that seem not to take stress (or perhaps other morphology attracts it). This impressionistic analysis of stress placement, based on my perception of the stressed syllable in each unit, would benefit from a systematic investigation of the acoustic correlates of stress. At the moment, however, it appears that word-level prominence targets syllables, regardless of how exactly that prominence is realized.

Also in his discussion of syllables as organizing units, Hyman (2010) argues that syllables may not be equally relevant in every language, even if they are universal. Hyman previously argued that Gokana (Ogoni) did not have syllables. Here, he argues against this earlier analysis and shows that stems in Gokana, which are CVV(C)VV, do indicate some marginal evidence for syllable structure. He proposes that, just as some languages vary in their attention to tone, stress, nasalization, or glottalization, there may be languages which attend very little to syllables and others, like English or Japanese, which more robustly rely on the syllable as an organizing structure. The answer to this question in Angaité remains to be seen, though this analysis suggests that Angaité attends to syllables at least enough to limit the number of segments that can appear in each one.

3.2 Syllable Structure Markedness

If we accept based on consonant distribution and stress assignment that Angaité has syllables, the next question is whether the syllables that appear are typologically marked or unmarked. A large portion of the literature on syllables is devoted to defining syllable structure markedness. One key concern is syllable weight and which parts of the syllable contribute to it. In general, Hyman (2006) notes that onsets do not contribute to syllable weight. Syllable rimes contribute to weight; a diphthong creates a heavy syllable compared to a monophthong, and a closed syllable creates a heavy syllable compared to an open one. Whether different rime configurations (diphthongs vs. long vowels vs. coda consonants)

contribute equally to syllable weight is a language specific question.

Often, syllables prefer to have an onset consonant, ideally just one, and an open coda position, making CV a common syllable structure (Carlisle 2001). The preference for single-consonant onsets is so strong that it is sometimes formulated as an implication: if a language has syllables with n onset consonants, the language will also have syllables with $n-1$ onset consonants. This implication holds unless $n = 1$; CV syllables do not imply the presence of headless (V) syllables. Diachronic evidence supports this preference for CV syllables. Over time, CCV syllables tend to reduce the number of onset consonants, and CVC syllables tend to lose coda consonants, moving toward CV syllables (Carlisle 2001).

Though cross-linguistic and diachronic evidence suggests that headed syllables are common and perhaps preferable to headless structures, exceptions do exist. For example, Breen and Pensalfini (1999) present evidence from reduplication and syllable-reordering games (Rabbit Talk) to show that Arrernte (Arandic, Central Australia) includes only underlying VC structures. Consonant-initial and vowel-final words surface only due to phonological processes related to vowel weakening and deletion. However, in the case of Angaité, the data do not currently support a VC-only (or VC-preferable) analysis. Most strikingly, all words in Angaité begin with consonants; in Arrernte words that begin with consonants are shown to have an underlying weak /e/ which deletes in phrase-initial position. There is no evidence for a similar process in Angaité. Furthermore, unlike in the Arrernte case, in which Breen and Pensalfini (1999) argue that final vowels belong to the first syllable of the following word, in Angaité most words also end with consonants, making an underlying VC pattern unparsimonious if not impossible. Thus, though Arrernte's VC pattern provides an interesting counter-example to the implicational universal formulated in Carlisle (2001), a more economical analysis in Angaité is one which prefers headed syllables to headless ones.

Syllable structure is also sensitive to segment sonority. The Sonority Sequencing Principle (SSP) states that the most sonorous element in a syllable is a nucleus with sonority decreasing in either direction. Even languages like Polish with large tautosyllabic consonant clusters can be shown to attend at least minimally to the SSP in syllable structure (Bethin 2011). The SSP is not absolute, with two main violations found cross-linguistically: plateaus

and reversals. Plateaus, where the sonority does not change from one segment to the other, are more common than reversals, where sonority actually dips between the margin and the nucleus of the syllable (Carlisle 2001). The sonority hierarchy itself may be slightly language specific, leading to cross-linguistic variations in the SSP's realization.

Angaité syllables also obey the SSP. The most sonorant element in each syllable works as the nucleus (always a vowel), with sonority decreasing toward the margin. Angaité does not generally allow consonsonant clusters, but the one exception follows the SSP as well, with a glide closer to the nucleus followed by an obstruent.

3.3 Typology of Diphthongs

The main issue at hand in Angaité is the question of whether diphthongs are present. Sánchez Miret (1998) provides a typology of diphthongs that helps locate the patterns in Angaité in a cross-linguistic context. Diphthongs can be divided into two major groups: falling diphthongs where the gliding portion follows the steady-state nucleus, and rising diphthongs where the gliding portion precedes the nucleus. Falling diphthongs are cross-linguistically more common; Sánchez Miret (1998) proposes that this pattern appears because falling diphthongs make the nucleus more prominent while rising diphthongs do the opposite. An additional distinction between falling and rising diphthongs appears in Cangnan Southern Min Chinese, where Hu and Ge (2016) find that falling diphthongs are realized with one articulatory gesture while rising diphthongs are two more separate gestures. They propose that the pronunciation of falling and rising diphthongs may be language specific.

Sánchez Miret (1998) also highlights diachronic trends. For example, he notes that falling diphthongs tend to evolve into VC sequences while rising diphthongs tend to become CV sequences. That is, the gliding portion tends eventually to be realized as a consonant. Falling diphthongs also regularly evolve into rising diphthongs, which, in turn, often become monophthongs, while the reverse is rare; rising diphthongs do not often become falling ones.

Height distinctions also affect the realization of diphthongs and help describe how diphthongization articulates with vowel hiatus. Crosslinguistic evidence points to the ideal diphthong being one whose two elements have a maximally different sonority (determined

by height). In cases with adjacent vowels of about the same sonority languages prefer hiatus to diphthongization (Sánchez Miret 1998). So, sequences of /eɔ/ or /oɛ/, for example, are more likely to be realized as hiatus, while /aɪ/ or /oɪ/ more often become diphthongs. Since /i, u/ are the lowest sonority vowels, they are ideal candidates for diphthong offglides, combining with most other vowels without resulting in hiatus. Since diphthongs prefer their two elements to be maximally different, Sánchez Miret (1998) finds that /aɪ/ is the sequence most likely to be realized as a diphthong.

The VG sequences at issue in Angaité are predicted based on Sánchez Miret (1998). They consist of a vowel, either mid or low, followed by a glide, a low sonority element. The most common combination is, indeed, /aj/, where the two elements are maximally different in sonority. Per Sánchez Miret, these sequences are likely to be treated as diphthongs, though based on the distribution of consonants and syllable structures I argue that this is not the best analysis for Angaité.

3.4 Vowel Hiatus Resolution Strategies

As hinted in §3.3, diphthongs and vowel hiatus are closely related, where two juxtaposed vowels with highly different sonority are likely to be realized as diphthongs while vowels with similar sonority are more often realized as hiatus. Vowel hiatus can be resolved by forming a diphthong—by realizing one vowel as a glide or inserting a glide between them—or deleting one of the two vowels. Transition timing is likely a helpful acoustic cue for determining whether a segment is realized as a vowel or as a glide. For example, Jaggers (2018) finds that transition earliness in American English is the best indicator of whether a sequence begins with [j] or [i]; that is, whether hiatus has been resolved via glide formation.

Other acoustic studies of American English have found that hiatus resolution is sensitive not only to typological constraints but also to word (and morpheme) boundaries and other prosodic processes. Davidson and Erker (2014) examine three types of sequences: VGV sequences within a word, VV sequences within a word, and VV sequences across a word boundary (VBV). They find that, contrary to previous studies which had claimed that English always resolves hiatus with an epenthetic glide, in the VBV environment speakers

insert a glottal stop or glottalization when the following syllable is stressed. That is, they find that no glide appears when the hiatus spans a word boundary. They do, however, see glides resolving hiatus word-internally. Additionally, their sample shows that underlying VGV sequences with lexical glides are longer than hiatus resolved with epenthetic glides. Therefore, they suggest that there may be acoustic differences in glide realization depending on whether the segment is underlying.

These findings suggest several considerations for syllable structure in Angaité. First, they suggest that hiatus resolution strategies, if they occur, are likely sensitive to morpheme or word boundaries, and they indicate that lexical and derived glides may be realized differently, which would allow acoustic data to aid in determining whether a glide is underlying. Davidson and Erker (2014) also indicate that hiatus resolution strategies are sensitive to suprasegmental prosodic processes like stress assignment, which, in American English, contributes to predicting whether VBV hiatus will be resolved with a glottal stop.

An additional consideration that comes out of the literature is the distribution of diphthongs and vowel hiatus. Davidson and Erker (2014) note that languages which allow vowel hiatus often do not include diphthongs. Angaité, which does show evidence of hiatus (37), is therefore not predicted to use diphthongs, though, as shown by Sánchez Miret (1998), the VG sequences are prime candidates for diphthongization.

(37) *maek*
maek
'hunger' (20190722, 2, 1)

Vowel hiatus is not preferred in Angaité (it is uncommon), perhaps because hiatus necessarily creates headless syllables which, as discussed in §3.2, are more marked than syllables with onsets. However, words like (37), as discussed in §2.5.1, do seem to function like other bisyllabic words, particularly with respect to pitch, which tends to rise in final syllables. This behavior suggests that Angaité does allow vowel hiatus rather than resolving CVVC sequences into a single syllable with a diphthong. If future research finds both diphthongs and vowel hiatus in Angaité we might appeal to discussions of diachronic processes that affect diphthongs to help explain their co-occurrence.

3.5 Vowel-Glide Sequences in Angaité

This analysis of Angaité syllable structure considers a subset of the elicited corpus containing 624 individual lexical items: 357 from the first seven hours of elicitation and 267 from the last five hours. For this subset, each item was marked for syllable structure twice; once on an analysis that takes vowel-glide (VG) sequences to be diphthongs and one that takes VG sequences to be a vowel-consonant sequence (VC). The number of each type syllable in each token was recorded in a spreadsheet which provided the total number of each type of syllable on each analysis. I ultimately conclude that VG sequences are best analyzed as VC sequences rather than as diphthongs based on the fact this analysis produces fewer irregularities in Angaité's syllable inventory and consonant distribution.

3.5.1 Option 1: Diphthongs

With the sample corpus syllabified on the analysis that all VG sequences are diphthongs, the total number of syllables considered is 2,045. These syllables fall into four types: CVC, CV, VC, and V. On this analysis, then, the minimal syllable in Angaité is a V, which can have up to one consonant on either side. The number of syllables in each category on this analysis appear in Table 3.1.

Syllable structure	Count
CVC	1,027
CV	875
VC	100
V	43

Table 3.1: Total number of each syllable type with diphthongs

CVC is slightly more common than CV, with VC and V appearing much less frequently. This pattern is predicted by typological work suggesting that headless syllables are highly marked. This analysis also indicates that Angaité prefers closed syllables to open ones; VC is more common than V and CVC more common than CV. This preference for coda consonants runs contrary to predictions from typological literature, since codas are more marked than open syllables, though perhaps not substantially so.

Syllable weight in general may turn out to be more important than whether coda consonants appear. Angaité also has been argued to make use of a vowel-length contrast (see Unruh and Kalisch (2003) who discuss vowel length in Enlhet-Enenlhet languages). It may be that open syllables tend to have long vowels, creating a heavy syllable (CVV). On this analysis where VG sequences are diphthongs, many open syllables, though not all, contain a diphthong, which, again, creates a heavy syllable. Perhaps, then, Angaité prefers heavy syllables, either those with a long vowel, a diphthong, or a coda consonant.

However, treating VG sequences as diphthongs causes two notable irregularities, one related to syllable structures themselves and one related to consonant distribution. First, analyzing VG sequences as diphthongs introduces a syllable structure, V, which does not appear when VG sequences are analyzed as VC sequences (see §3.5.2). Furthermore, the distribution of the V syllables is highly restricted, appearing only after open syllables whose nuclei are diphthongs. The other three syllable types—CV, CVC, and VC—do not have such restrictions. CVC, CV, syllables can be word-initial or word-final. VC syllables do not appear word-initially, but otherwise appear after open syllables with either diphthong or monophthong nuclei. These other syllable templates and their restrictions hold when VG sequences are analyzed as VC sequences, while V syllables appear only when VG sequences are analyzed as a diphthong.

Furthermore, while glides pattern like all other consonants in onset position, treating them as diphthong offglides following vowels means that they never appear as codas. Since all other consonants can be codas, this behavior for glides is not predicted. Treating VG sequences as diphthongs, then, creates two substantial irregularities: a V-only syllable with a restricted distribution and a reduced distribution of glides compared to other consonants.

3.5.2 Option 2: Consonantal glides

Alternatively, the sample corpus can be syllabified on the analysis that all VG sequences are VC sequences with the glide forming part of the rime with the preceding vowel or the onset of the next syllable. In this case, the total number of syllables is 2,044. These fall into three types: CV, CVC, and VC. The total number of each of the three syllable

types appears in Table 3.2. Here, CVC syllables are much more common than CV ones, and

Syllable structure	Count
CVC	1,194
CV	823
VC	23

Table 3.2: Total number of each syllable type with VC analysis

VC syllables are more infrequent than on the diphthong analysis. The latter observation fits with the typological literature showing that headless syllables are highly marked. The fact that CVC syllables are most common when CV is also available is perhaps surprising from a typological standpoint but consistent with the pattern in §3.5.1. Regardless of how VG sequences are analyzed, then, Angaité includes a high number of CVC syllables, perhaps again indicating that heavy syllables are generally preferred.

However, this analysis also creates irregularities in syllable structure and consonant distribution. First, it introduces a CVCC syllable structure which appears once, (38).

(38) *'a.nayk.mas.ka.ma'*
ʔanajkmaskamaʔ
'You/she sing(s)' (20190719, 1, 42)

Example (38) appears consistently in verb constructions with 'sing', making this verb the only one in this corpus with an irregularly shaped stem. Further morphological parsing of these forms will hopefully suggest either an explanation based on affixation, or a diachronic one. If VGs are VC sequences, then glides are the only consonants that create tautosyllabic consonant clusters. These clusters do obey the SSP, with the glides appearing closest to the nucleus and the less-sonorous consonant closer to the margin, so they are not typologically uncommon in this respect, though within the language they are exceptional.

This syllabification also results in some oddiities which can probably be attributed to quick speech rather than the underlying representation. The corpus includes one example pronounced as in (39), where a palatal nasal surfaces with no clear following /j/.

(39) *nyavakha'*
[ɲawakhaʔ]
**njawakhaʔ
'our town' (20190708, 2, 66, EG)

However, unlike the examples in (38), this word is attested elsewhere in the corpus, where it is pronounced as in (40).

(40) *nengavakha'*
[ne.ŋa.wak.haʔ]
neŋ-awakhaʔ
1PL.POS-town
'our town' (20190723, 2, 24, EF)

The underlying form in (40) is consistent with my current analysis of the possessive prefixes in Angaité, making it preferable to an analysis which includes a consonant cluster and a glide-initial stem. Taken together, (39) and (40) suggest that some surface consonant clusters can be attributed to fast speech rather than irregular underlying representations.

The corpus also includes one example of a V-only syllable, in (41).

(41) *ja.ma me.le.o.ke'*
jama meleokeʔ
'spring' (20190722, 1, 9, EG)

Given that EG frequently omits consonants (particularly glides) and, occasionally, syllables, I suspect that this example is probably also a result of casual speech and does not fully reflect the underlying form. A more robust morphological analysis will illuminate whether all unexpected forms result from the rapid pronunciation of underlying morphemes with predicted shapes. Together with (38) and (40), the total number of irregular forms when VG sequences are treated as VCs is three, compared to treating VG sequences as diphthongs, which introduces an additional, predictably-distributed syllable type and an asymmetrical distribution of glides.

3.5.3 Evaluation of syllable structure

Based on this corpus, analyzing VG sequences as VC sequences, where the glide is syllabified either as a coda (followed by another consonant) or an onset (followed by a vowel) is the better analysis for Angaité. It creates fewer irregularities, with essentially one verb stem and two other anomalous forms to account for. In general, glides pattern like other consonants, appearing in onset and coda position at both word-margins and word-internally. Taking glides to be diphthongs introduces two additional difficulties: V-only syllables with a highly restricted distribution and glides which pattern differently from other consonants by never appearing in coda position.

A further point in favor of glides being treated as consonants in all cases comes from typological work that indicates that languages which allow vowel hiatus do not often include diphthongs. Both analyses do include instances of vowel hiatus, meaning that analyzing VG sequences as diphthongs not only creates more inconsistencies within Angaité but also posits a cross-linguistically unusual pattern. Taking glides as consonants conforms better to generalizations about how diphthongs and vowel hiatus pattern, making this option more parsimonious on both language-internal and cross-linguistic parameters.

While this look at syllable structure provides a start to describing the system, it still leaves open various avenues for future research. First, and perhaps most obviously, this analysis has assumed that *either* VG sequences are diphthongs *or* they are vowel-consonant sequences. It does not consider the possibility that the language might include both diphthongs and VGs as VC sequences. I approached the problem in this way to avoid opportunistically describing glides as consonantal when they did not create a consonant cluster and as diphthong-offglides when they did. Further examination, however, may suggest that in some cases VG sequences act more like a single unit (diphthongs) and in other cases the glide is more consonant-like.

An additional, and related issue, is whether affixation affects syllable structure. For example, the odd cluster in (38) may be the result of affixing T/A/M morphology on the verb stem. Further description of Angaité morphology should clarify this issue by indicating whether word-internal morphological boundaries can cause irregular consonant clusters—in

which case we would anticipate finding them in other stems—or whether the 'sing' stem is unique. Further morphological analysis can also clarify whether VG sequences act more like one unit or two; for example, a VG sequence that falls over a morpheme boundary or can be separated by another morpheme would present additional evidence for analyzing glides in these cases to be consonantal and not part of a complex nucleus. A morphological analysis may also enrich the description of syllable type distribution. For example, Gomes (2013) argues that VC syllables only appear as prefixes and that CVC syllables always appear word-finally. Based on the corpus, these particular generalizations do not hold across the board for Angaité, but a closer look at morphology may indicate that particular syllable structures are always paired with the same morphological classes.

This analysis also does not posit syllables with complex nuclei that have a CVVC form. Positing a CVVC structure would eliminate headless syllables entirely from Angaité. Such an analysis would also perhaps allow a unified treatment of long vowels (which here I treated as a single V) and vowel hiatus like what surfaces in (37). However, this strategy is not taken here. First, pitch and length in CVVC sequences appears to function more like bisyllabic sequences than a single unit. Additionally, stress assignment should treat CVVC's as either one syllable or two; determining whether this happens will require examining the acoustic correlates of stress in word-final CVVC sequences and investigating the possibility of metrical stress assignment. A CVVC syllable might restrict which vowels can combine in the nucleus, and in what order, while a two-syllable analysis predicts free combination. In the current dataset, we see most, though not all, logically possible vowel combinations, (42), with relevant segments **bolded**.

(42) a. *nempak**meas**ma'*
nem-pakmeasma?
1PL.DIR-speak
'We speak' (20190708, 3, 61)

b. *kelvane**maet***
kelwanemaet
'wolf fish' (20190708, 1, 45)

c. *'apke**nao'***

ʔapkenaoʔ

'man' (20190708, 1, 28)

d. *sose**khoe'***

sosekhoeʔ

'morning' (20190708, 1, 10)

Thus, the data so far, with a near-complete set of possible vowel combinations in CVVC sequences, which also share suprasegmental features with bisyllabic structures, point to CVVC strings as better analyzed as two syllables rather than one.

In addition to disambiguating the issues discussed above, this analysis will eventually benefit from a controlled acoustic study like those done in Hu and Ge (2016), Jaggers (2018), or Rivera Castillo (2017), which characterize vowels and glides based on duration, intensity, and articulatory movement. An analysis of this type would help situate Angaité not only within literature about the typology of phonological systems but would also contribute to a more robust cross-linguistic description of the phonetic characteristics of these segments, whose articulatory and acoustic properties are still not fully understood.

Chapter 4

Vowels

This section briefly describes what can currently be said about vowels and how this system articulates with work on related languages. Along with my transcriptions of each vowel, I performed a brief quantitative analysis of a subset of the corpus. I measured beginning and end pitch; pitch minimum and maximum; intensity, F1, F2, and F3 at the midpoint; and duration of each vowel in a set of 73 words, resulting in a total number of 430 measured vowels.[1] Of these tokens, approximately half were spoken by each consultant (227 from EG and 203 from EF). Words ranged from two to five syllables, with 20 two, three, and four-syllable words, and 13 five-syllable ones. The vowels in each word were marked in a text grid in Praat; the start and end points of F2 and F3 determined the start and end points of each vowel. Vowels were labeled with an IPA symbol (based on perceived vowel quality) and a number that corresponded to the interval in the text grid (first vowel in a file was labeled 1, the second, 2, etc.).

4.1 Number of Phonemes

Like other Enlhet-Enenlhet languages, Angaité has three phonemic vowels. These phonemes have been described in other languages (for example, Sanapaná, by Gomes (2009; 2013), and van Gysel (2017)) as /a, e, o/. Plotting the formants of all the measured tokens confirms this analysis, showing three clusters, Figure 4.1.

These data are messy—measurements were taken from vowels in every syllable of each word and were not controlled for the surrounding consonant, which certainly impacts vowel articulation. Even so, the tokens cluster into three groups which correspond to the three

[1] Measurements were taken and recorded using a Praat script from Elizabeth Wood designed to measure vowel formants and duration. I modified this script to include measurements of intensity and pitch.

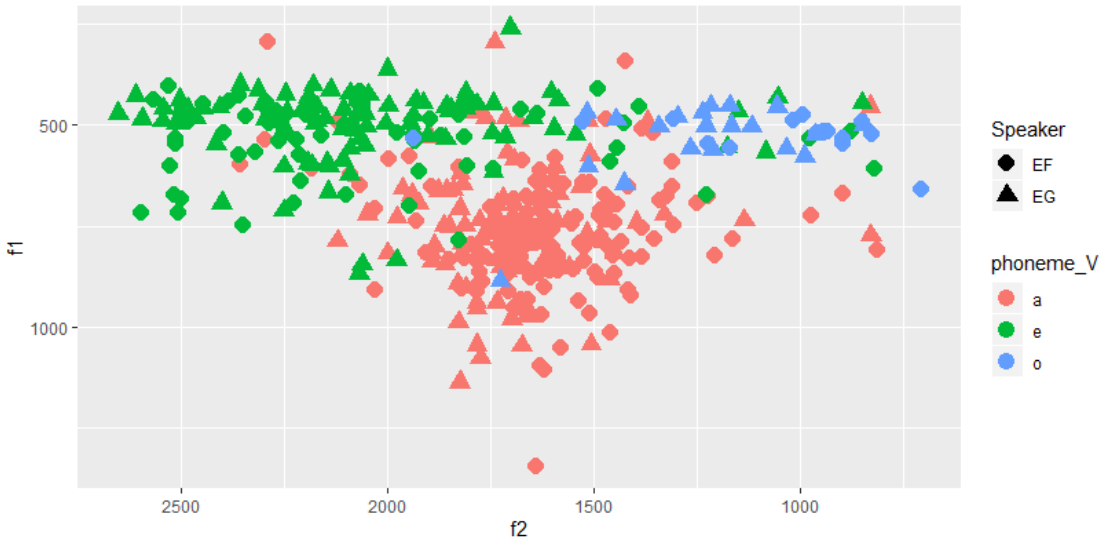

Figure 4.1: F1 and F2 values for Angaité vowels

anticipated phonemes. This pattern is even more notable when only the non-low vowels are examined, as in Figure 4.2, which shows only the vowels that were coded as [e, ɛ, ɪ, i] in red and [o, u] in blue. This graph shows relatively little overlap between the two categories. Furthermore, this first pass suggests that a three-phoneme analysis is a good fit for Angaité, since these tokens do not show notable clusters within the non-low front and non-low back tokens. While some vowels are higher than others, separate clusters that might indicate a meaningful category difference do not emerge.

However, a system with no phonemic high vowels is typologically uncommon compared to a three-vowel system with /i, u, a/ or /i, o, a/. If a more typologically common analysis can be adopted, I would prefer to use it so as not to present Angaité as overly uncommon compared to other languages with comparably small vowel inventories. A close examination of the distribution of allophones in Angaité can either convincingly argue for /e/ and /o/ as the phonemic representation of the non-low vowels or, alternatively, suggest that /i/ and /u/ can equally well be analyzed as the underlying phonemes.

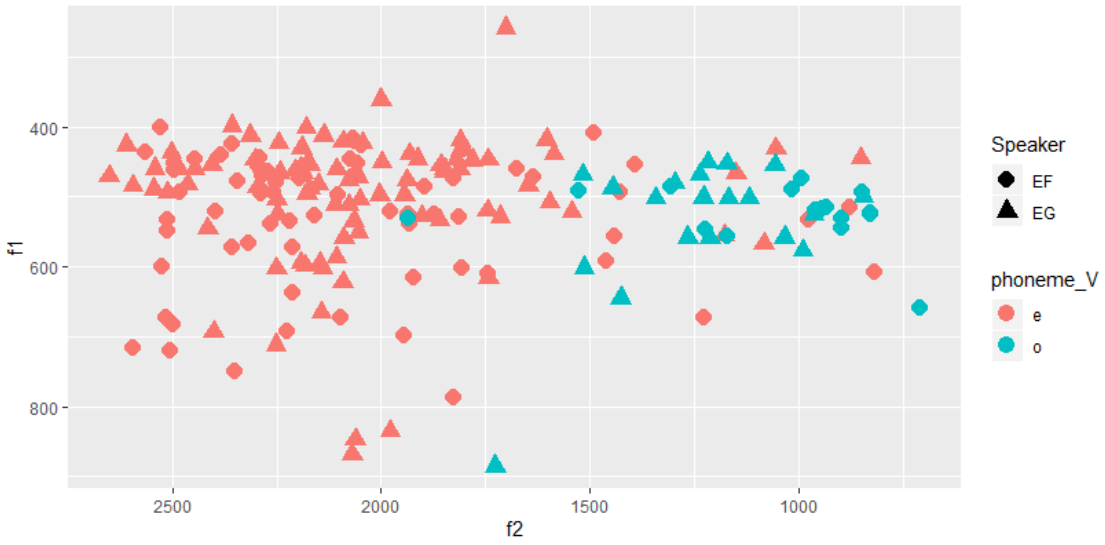

Figure 4.2: F1 and F2 of non-low vowels in Angaité

4.2 Non-low Vowel Allophones

Gomes (2009) suggests that the non-low front phoneme /e/ in Sanapaná has three allophones: [i, e, ɛ]. He proposes that these phones are phonologically-conditioned allophones of a single phoneme because substituting one for the other, though it does not produce a new word, results in speaker correction. Similarly, he describes two allophones of the non-low back phoneme /o/: [o, ɔ]. Gomes does not describe the distribution of these allophones, nor does he explicitly justify his choice of /e/ and /o/ as the underlying phonemes. He also describes three long vowels, [eː, oː, aː]; the fact that other allophones do not have long variants may implicitly underlie the choice of phonemic /e, o/.

When marking vowels for measurement, I transcribed four non-low front allophones: [e, ɛ, i, ɪ] and two non-low back phones. Unlike what Gomes reports for Sanapaná, I noted the back vowels to be [o] or [u], not [ɔ]. The back vowels are less frequent in Angaité; perhaps an [ɔ] allophone would surface in a larger corpus. I also transcribed [ɪ], which Gomes does not. In fact, this allophone is more frequent than [i] in subset used for vowel analysis, with 21 appearances compared to only 6 of [i]. One potential explanation for this difference is that front vowels in Angaité are slightly lower or more central than Sanapaná, leading to more instances of [ɪ] (and [ɛ]) than Gomes finds.

4.3 Distribution of Allophones

No work on the distribution of vowel allophones has been done for Enhlet-Enenlhet languages. Impressionistically, /e/ often appears to be lower and more central in what I have marked as unstressed syllables, appearing as /ɛ/. Higher allophones also tend to appear adjacent to velar sounds, particularly /k/. Since velar sounds are often characterized as [+high], systematic correspondences of higher allophones in the presence of /k/ throughout the corpus may indicate that velar sounds condition vowel raising.

Thorough examination of the distribution of these allophones will also provide a better justification for the selection of /e/ and /o/ as the basic allophones of the non-low phonemes. The selection of a basic allophone of the phoneme makes little difference in how the vowel system works within Angaité: regardless of which is selected, the language has a two non-low vowels, one front and one back. However, as noted above, a closer examination of the distribution of these phones and the articulatory targets that underlie them will confirm whether the underlying representations of the non-low categories are better classified as /i, u/, /i, o/, or /e, o/, contributing to Angaité's broader typological contextualization.

4.4 Vowel Length

According to Unruh and Kalisch (2003), Enlhet-Enenlhet languages, with the exception of Toba-Enenlhet, contrast long and short vowels. As noted above, Gomes (2013) argues that /e, o, a/ have long variants in Sanapaná. In contrast, van Gysel (2017) prefers not to analyze vowel length as contrastive due to a lack of minimal pairs contrasting only on the basis of vowel length.

In the sample of vowels that I measured in Angaité, no significant length differences appear. Figure 4.3 shows the length of each token, grouped by phoneme. However, these measurements do not present a complete picture. Vowels from every position in the word were measured, which may obscure nuanced length differences that depend on position or stress. Furthermore, within words, some vowels are certainly longer than others, often the penultimate vowel in the word. For example, take /ʔaaɬaʔ/, represented in the spectrogram in Figure 4.4, with vowels marked in a Praat text grid. The first vowel (*a1*) measures

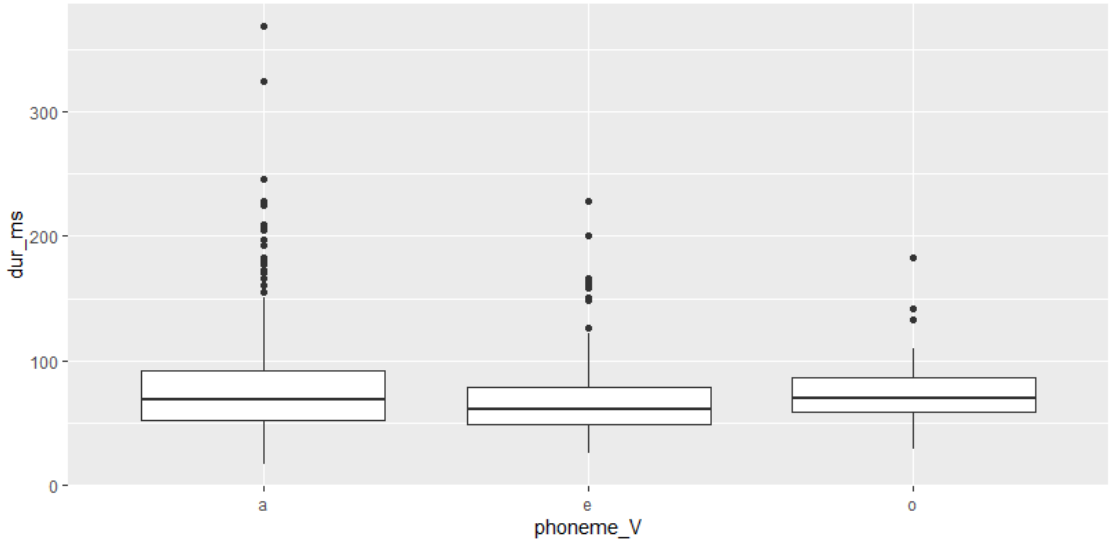

Figure 4.3: Length of vowels in Angaité (in ms)

224.867ms, and the length of the second (*a2*) is 73.986ms. These measurements show that *a1* is three times longer than *a2*. In words that I transcribed with a long vowel, the vowel that was perceived to be long was usually at least 1.5 (and sometimes up to 4) times longer than other vowels in the word. This first look at vowel length suggests that, while a compilation of all vowels in a sample does not show significant length distinctions, a controlled sample accounting for position, speaking speed, and stress, may reveal a distinction. As noted in Chapter 3, vowel duration may also be related to syllable structure; possible analyses of long vowels are as CV.VC(C) sequences over a syllable boundary or as CVV(C) syllables with complex nuclei. The third alternative (adopted here) considers long vowels to be tautosyllabic CV:(C) segments. Whether a length distinction is predictable based on other factors—stress, position, syllable type—is an additional puzzle that can only be examined alongside a reliable analysis of word boundaries, morpheme boundaries, and stress assignment.

4.5 Vowel alternations

Non-low vowels, particularly /o/, sometimes alternate in Angaité. For example, my dataset includes one variable token, 'river', which was pronounced by EF two different ways

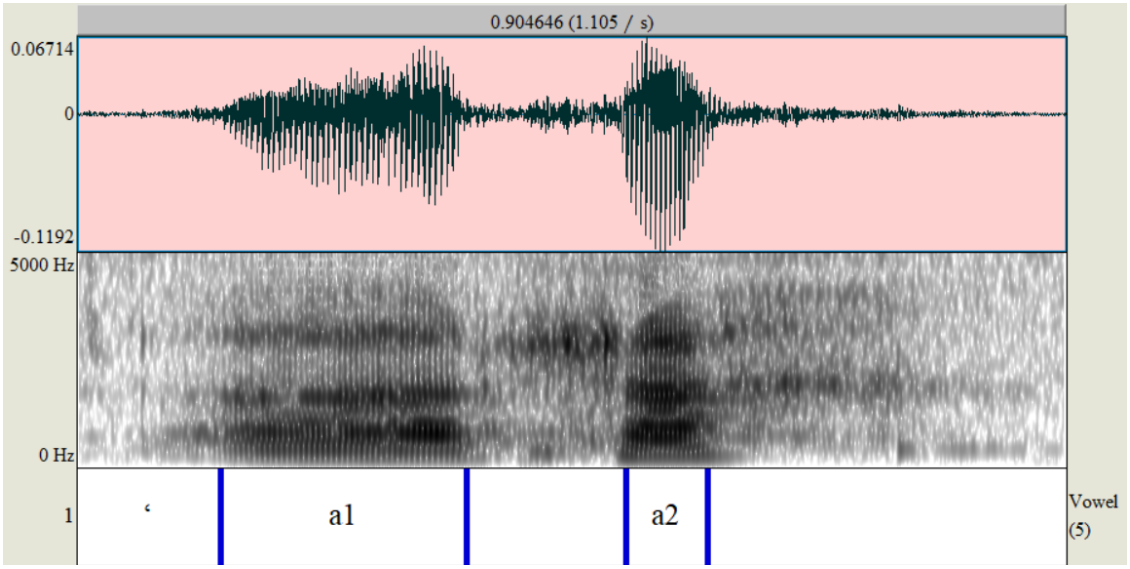

Figure 4.4: Spectrogram for *'aalha'* in Angaité produced by EF

on different days, (43).

(43) a. *vatsom*
[watsom]
‘river’ (20190722, 1, 5)

b. *vatsam*
[watsam]
‘river’ (20190710, 1, 8)

Formant listings from the middle of these two vowels (start and end points of the vowels determined by beginning and end points of F2) indicate substantial differences, presented in Table 4.1. The F1 values listed in Table 4.1 indicate that the vowel I have transcribed as [o]

	[watsom]	[watsam]
F1	669.6237	828.1520
F2	2839.6367	1549.4140
F3	3363.0106	2654.5972

Table 4.1: Mid-point formants for final vowel in ‘river’ as pronounced by EF

is substantially higher than the one transcribed as [a], but F2 values are more inconclusive,

suggesting, surprisingly, that the [o] vowel is more front than the [a] one. As noted above, the non-low vowels in general have a more variable pronunciation, which may account for some of this discrepancy.

The alternation in *vatsom* exemplify a larger pattern of vowel harmony in the language. Alternations in Angaité are not limited to /o/ ∼ /a/; variation is also particularly notable in *sosenhe'*, 'net bag', which is also alternately pronounced as *sesenhe'* by EG. In both these cases, one of the two pronunciations has identical vowels across the whole word. Many other examples of words with identical vowels appear in the corpus, though alternating pronunciations are not common. See Table 4.2 for examples cited elsewhere in this paper.

Angaité		English	Source
yaapa'	jaapaʔ	'money'	(20190723, 1, 59)
'aktaykamaha	ʔaktajkamaha	'You/she works'	(20190720, 2, 14)
'apna'at	ʔapnaʔat	'your/his face'	(20190722, 2, 50)
makva'	makwaʔ	'peanuts'	(20190722, 2, 16)
naata'	naataʔ	'bird'	(20190722, 1, 11)
talhnaam	taɬnaam	'afternoon'	(20190708, 1, 7)
nepkesek	nepkesek	'sheep'	(20190716, 1, 57)
metke'	metkeʔ	'not, none'	(20190822, 1, 14)
yemen	jemen	'water'	(20190708, 2, 97)
metekteng	metekteŋ	'duck'	(20190708, 1, 72)
pooko'	pookoʔ	'hot'	(20190710, 3, 50)
mo'ok	moʔok	'other'	(20190708, 3, 35)

Table 4.2: Examples of words with uniform vowels in Angaité

Perhaps unsurprisingly given the other patterns noted here, when vowels within a word are identical, /a/ is most commonly used, followed by /e/, with /o/ the least frequent. Though identical vowels do frequently occur, words with this characteristic are by no means the majority in Angaité. Whether this pattern results from an historical process of partial vowel harmony, synchronically active harmonization, or simply chance co-occurrence is still unclear. As is the case with determining the distribution and articulatory targets of the non-low allophones, this question will be best investigated in elicitation that tests morphological paradigms for predictable patterns.

4.6 Creaky voice and vowels

Many vowels in the corpus are creaky voiced. In some cases, this change in phonation is an effect from an adjacent glottal consonant, as is the case in *yephopay'a*, 'cloud'; the /aj?a/ sequence appears in Figure 4.5. Praat measures the local jitter in the vowel following the glottal stop to be 0.879%, compared to 0.028% in the preceding vowel. However, the

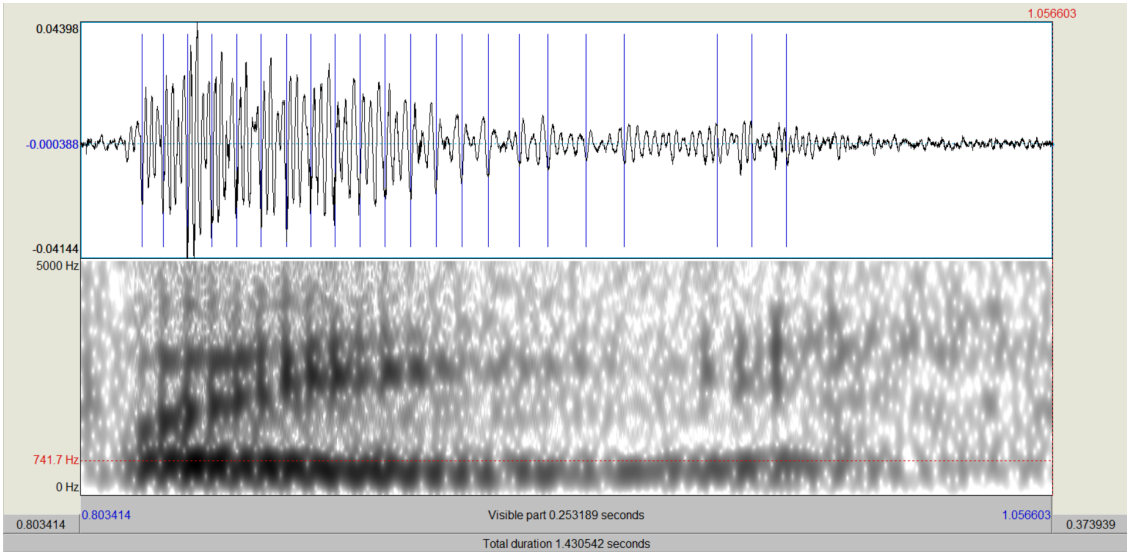

Figure 4.5: Spectrogram of /aj?a/ in *yephopay'a* (2019078, 2, 41, EF)

interaction between creaky voice and glottal stop, if any, is not always so clear. A subset of word-medial vowels—some but not all of the longer ones—appear to be geminate, with the second segment creaky-voiced. For example, the vowel in the the second syllable of *kemhaava'* 'puma' has a local jitter of 1.794%, accompanied by a falling pitch (Figure 4.6).

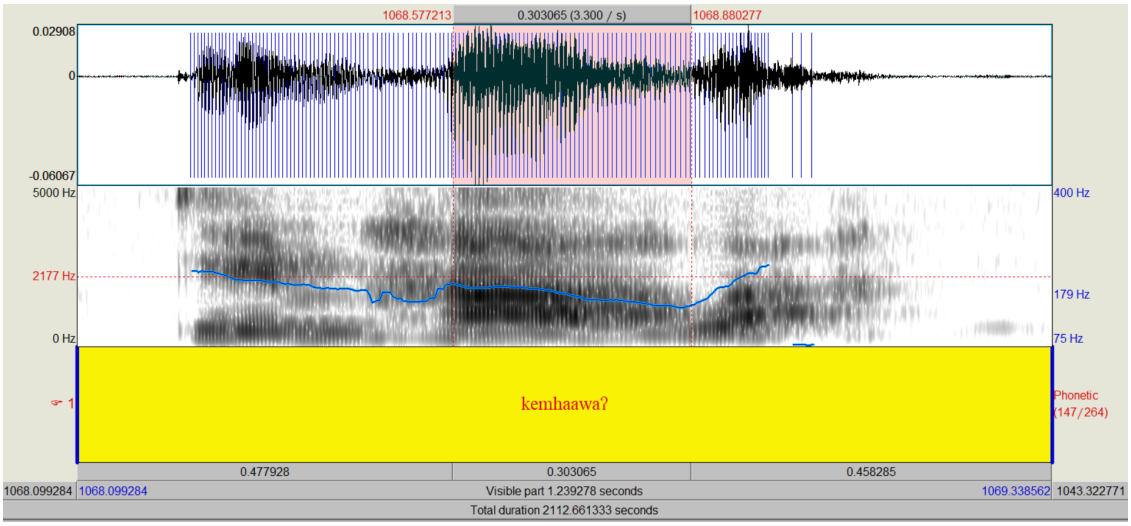

Figure 4.6: Spectrogram of *kemhaava* (20190708, 1, 46, EF)

More notable than the high level of jitter on the entire vowel, however, is that jitter in the second portion (from the beginning of the dip in pitch to the end) is 2.841%, and only 0.529% in the first half. Thus, jitter clearly delineates two distinct portions, one of which is creakier than the other. In Figure 4.6, the creak is accompanied by a falling pitch, but this is not always the case. Creak on the final vowel in *nepyayaam* (Figure 4.7) is accompanied by rising pitch. The vowel here is less creaky than in Figure 4.6, but the spectrogram nevertheless

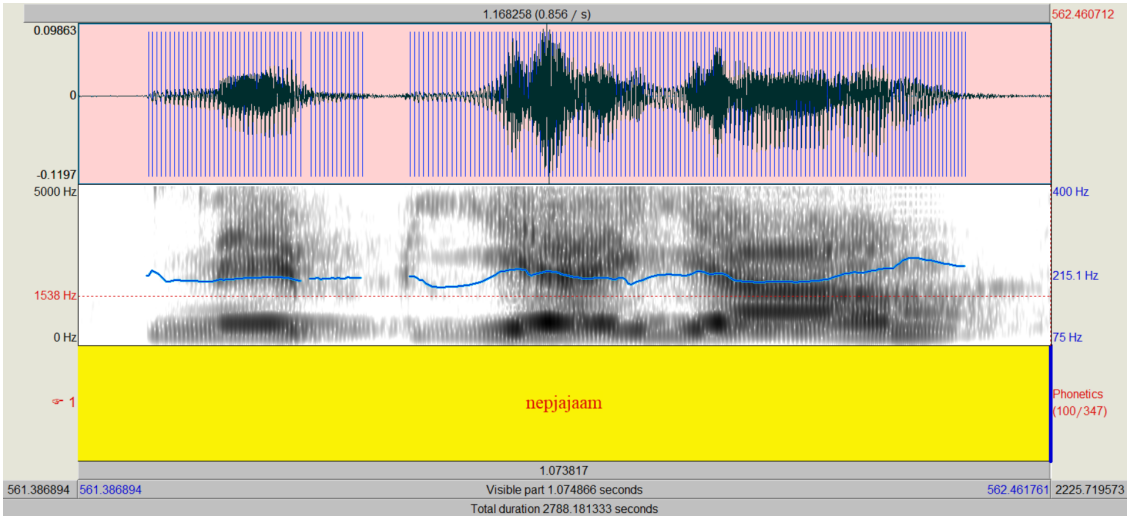

Figure 4.7: Spectrogram of *nepyayaam* (20190708, 2, 27, EF)

shows two distinct parts of the vowel, the second having lower intensity and higher jitter.

Creak does not always accompany pitch changes. For example, *heesevaske'* 'It loves me' shows a pitch rise in the second half of the /e/ in the first syllable (see Figure 4.8), but jitter is only 1.169%, lower than the jitter on the long vowel in *kemhaawa'*, 'puma'. The

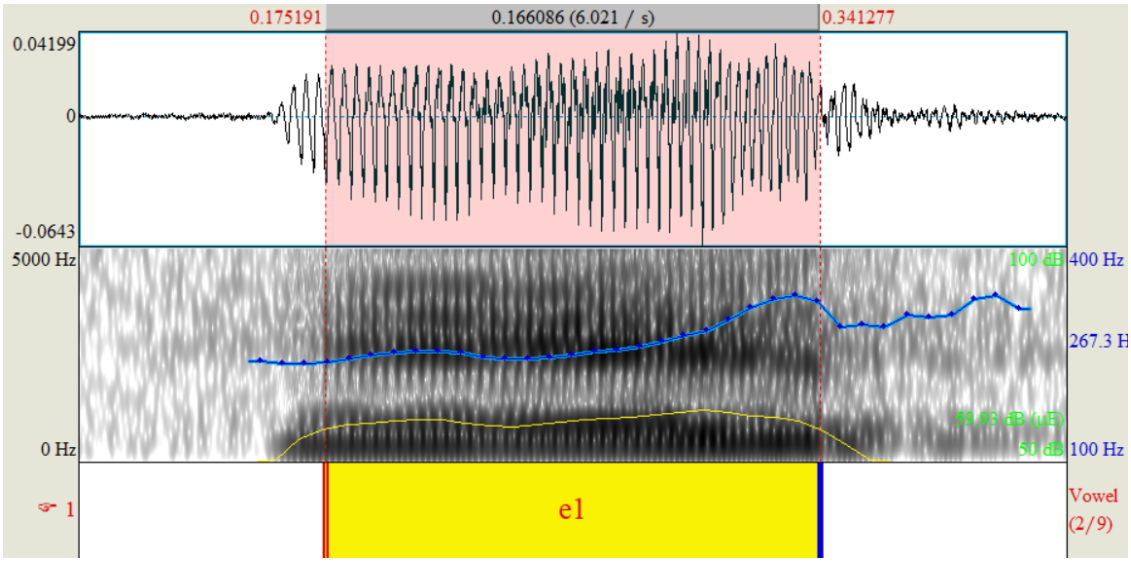

Figure 4.8: Spectrogram of first /e/ in *heesevaske'* (20190719, 1, 17)

beginning of the pitch rise is associated with a dip in amplitude (yellow line), indicating that in this case an underlying /ʔ/ may be realized as an intensity decrease and pitch increase on the vowel. Pitch also often dips and then rises toward the end of longer words, as in *kamelangkoho'* 'slow(ly)' (Figure 4.9) and *'apqaqheema'* 'You/he lives (in)' (Figure 4.10).

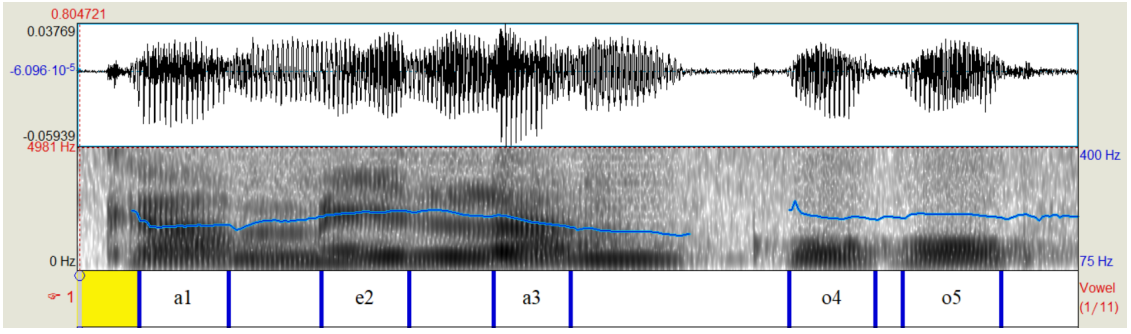

Figure 4.9: Spectrogram of *kamelangkoho'* (20190723, 2, 7, EF)

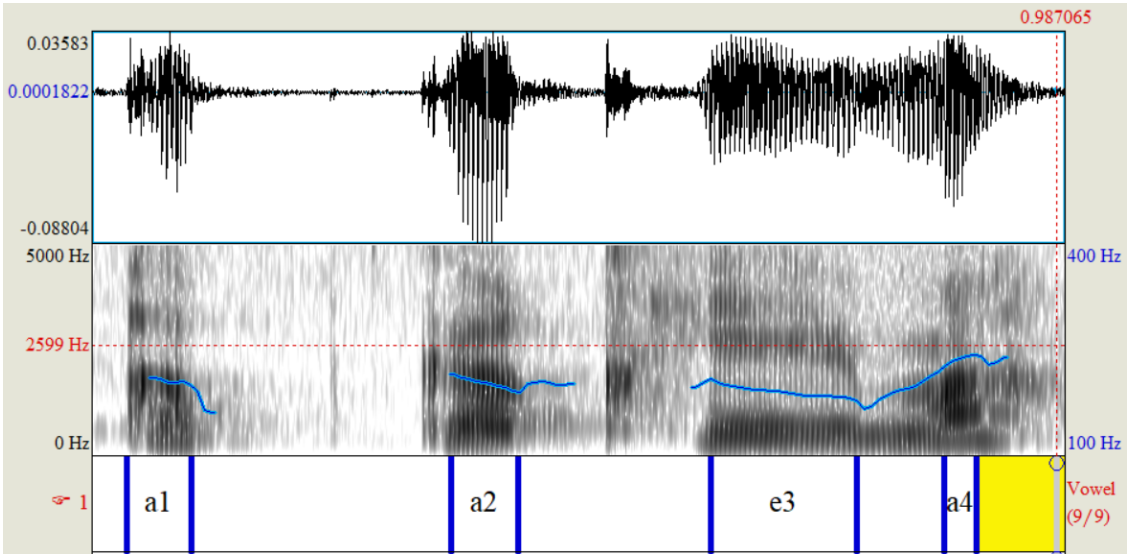

Figure 4.10: Spectrogram of *'apqaqheema'* (20190723, 2, 25, EG)

However, jitter on the vowels adjacent to this pitch valley is not particularly high (Table 4.3), indicating that a different process may underlie pitch dips and rises in cases like *heesevaske'* and *kamelankoho'* or *'apqaqheema'*.

Token	Pre-pitch dip	Post-pitch dip
kamelangkoho'	1.197%	1.871%
'apqaqheema'	0.608%	1.257%

Table 4.3: Jitter in vowels surrounding pitch valley in *kamelangkoho'* and *'apqaqheema'*

One explanatory hypothesis is that Angaité uses tone with a (perhaps) relatively low functional load and these words have a final high tone. Alternatively, pitch changes may reflect other prosodic processes. Because this corpus comprises primarily words in isolation, word and utterance-level prosodic phenomena may be conflated. Therefore, whether the interaction between creak and pitch is better understood as tone, a segmental feature, or an effect of a prosodic boundary is unclear. Future work might test words in carrier phrases and compare their suprasegmental features to words in isolation and in natural discourse. Regardless of the underlying cause(s), vowel creakiness is clearly rule governed: words pronounced with creak are generally consistently done so. Creak also appears with both speakers, suggesting that it is not an idiosyncratic feature of a single speaker.

Chapter 5

Conclusion and Future Research

Data for this study are drawn from data elicited from two speakers during my pilot trip in summer 2019. An eventual robust analysis of Angaité phonology will rely on a larger group of participants. Additionally, though an external microphone was used to try to limit background noise by placing the microphone as close to speakers as possible, my recording setup was imperfect. Substantial background noise and static still appear on my recordings, and speakers occasionally overlap. Future work will minimize this issue by utilizing individual head-mounted cardioid microphones during elicitation and individual narration sessions and only using a microphone with a broader range when recording in less-structured environments such as public events or large, group meetings. Though the data used here is flawed, they allow a step toward describing the phonological system of Angaité, which has not been detailed by other studies.

This paper has been primarily concerned with consonants and their distribution in Angaité, developing an inventory of 14 consonants in the language, organized into three possible syllable types, CV, CVC, and VC. The consonant inventory developed for Angaité in this paper is similar to the one proposed for Sanapaná by van Gysel (2017), with the addition of the voiceless uvular stop /q/. This similarity is unsurprising given the close relationship between the two languages, as indicated by Unruh and Kalisch (2003). The language includes a series of voiceless stops /p, t, k, q, ʔ/, a set of three nasals, /m, n, ŋ/, two fricatives, /s, h/, a lateral /l/, a lateral fricative /ɬ/, and two glides, /w, j/.

All consonants can appear in both syllable onset and coda position, and the majority can appear either word-initially or word-finally. The exceptions to this generalization are /ŋ/, which does not appear word-initially in these data; /q/ which does not appear word-finally; and /l/, which, based on borrowing accommodations from Spanish, is prohibited

word-finally. In Angaité, glides function like other consonants rather than in alternation with vowels, and they appear in both onset and coda position.

The minimal syllable in Angaité is CV, with CVC and VC options also available. VC syllables only appear in cases of vowel hiatus, which are relatively rare; headless syllables do not surface word-initially in this dataset. CV syllables only rarely appear word-finally. Gomes (2013) frames this pattern as a phonological process, arguing for an epenthetic word-final /ʔ/ to avoid final open syllables. At this point, my data do not provide evidence that final /ʔ/ are epenthetic in Angaité, and I analyze them as underlying. However, creak does seem to be an utterance final phenomenon in Angaité; words which utterance finally, and in isolation, ended with [ʔ] did not always do so medially. I have selected here the realization which appears most consistently over all instances of a given token. However, this tendency toward utterance-final creak means that words which were only elicited in isolation and end with /ʔ/ may have underlying final open syllables with the [ʔ] attributed to phrase level prosodic processes. This question is best left to future elicitation using carrier sentences. Perhaps because open syllables are rarely final, CVC syllables are most frequent, followed by CV, and then VC. Some minimal irregularities appear with this account of syllable structure; one form contains a consonant cluster and one a V-only syllable. However, this analysis, taking glides following vowels as consonantal, is preferable to one which takes VG sequences to be diphthongs. The diphthong alternative creates both a new syllable type (V) and complicates the distribution of glides, which would never appear in coda position.

Angaité's vowel inventory, though not thoroughly examined here, includes, as has been proposed for other Enlhet-Enenlhet languages, three phonemic vowels, which I take to be /e, o, a/. Vowels present several fruitful directions for future research. Other Enlhet-Enenlhet languages have been argued to include a length distinction which did not robustly surface in these pilot data, though some vowels are clearly longer than others. Whether this length is predictable (often seeming to appear before an apparently stressed syllable) and contrastive, is a question that can only be addressed alongside a thorough description of the acoustic correlates of stress and word and phrase-level processes that affect duration.

This analysis also does not thoroughly consider glottalization and creaky voice and how they interact with consonants. Glottalization, which appears on some vowels and some

word-final sonorants, is one possible realization of underlying phonemic glottals, /h, ʔ/. Pitch changes are also often associated with glottalization on vowels which, again, may be a surface realization of an underlying consonant. I have not analyzed all glottalization (and associated pitch rises) as underlying glottal stops here, but it may eventually turn out that at least some words with creaky voice do have an underlying /ʔ/. As discussed in the section on vowels, stress and pitch have also not been thoroughly addressed, though these features also interact both with individual segments and with larger units like words or phrases.

Additional morphological analysis will be extremely helpful in solidifying the analysis of syllable structure and working out which segments are underlying when the current data are inconclusive. For example, working out the morphemes that appear in verbal constructions, and then testing different permutations, should indicate whether creaky voice on vowels is best understood as an underlying /ʔ/. I predict that /ʔ/ should fully surface as an onset, so if creaky voice remains even when a vowel is consonant-adjacent, it is likely better analyzed as a prosodic feature and not the result of an underlying /ʔ/ onset. Morphological analysis should also help determine whether consonants at word margins are underlying or epenthetic, and whether segment length is predictable based on word position.

Eventual research will also work from spontaneously produced speech and a corpus of texts from a wider variety of speakers. This breadth of data will allow a better analysis of dialect variation and idiosyncratic pronunciations. Naturalistic discourse will be particularly helpful with morphological analysis which, as discussed above, also strengthens phonological analysis. Spontaneous discourse is useful for phonological analysis in its own right, as some suprasegmental features may manifest differently in running speech. Describing continuous or fast speech and comparing it to elicitation of isolated words or controlled carrier sentences will also help illustrate phenomena like deletion or coarticulation which are less prominent in carefully-produced words. The discussion here is a jumping off point for addressing these more interrelated morphological and phonological questions.

Appendices

Appendix A

Abbreviations

The following abbreviations are used in this paper:

1	first-person
2	second-person
3	third-person
A	agent (subject of transitive verb)
DIM	diminutive
DIR	direct
F	feminine
INV	inverse
M	masculine
O	object (of transitive verb)
PL	plural
POS	possessive
SG	singular

Appendix B

Scientific Name Glossary

'algarrobo negro'	(*Proposis nigra*)
'anaconda'	(*Eunectes murinus*)
'Azara's night monkey'	(*Aotus azarae*)
"Carandilla palm'	(*Trithrinax schizophylla*)
'collared peccary'	(*Tayassu tajacu*)
'Dutchman's pipe'	(*Aristolochia* genus)
'great-horned owl'	(*Bubo virginianus*)
'guinea hen'	(*Numida meleagris*)
'jaguar'	(*Panthera onca*)
'jungleplum'	(*Sideroxylon obtusifolium*)
'karaguatá'	(*Bromelia pinguin*)
'kururu pytã', Corraline frog	(*Lepodactylus laticepts*)
'maned wolf'	(*Chrysocyon brachyurus*)
'manioc'	(*Manihot esculenta*)
'palo santo'	(*Bursera graveolens*)
'passion flower'	(*Passiflora caerulea*)
'poroto del monte'	(*Capparis retusa Griseb*)
'puma'	(*Puma concolor*)
'rhea'	(*Rhea americana*)
'seven-banded armadillo'	(*Daypus septemcincus*)
'tagua, Chaco peccary'	(*Catagonus wagneri*)
'three-banded armadillo'	(*Tolypeutes* genus)
'timbo'	(*Albizia inundata*)
'toad'	(*Ceratophrys cranwelli*)
'tuca tuca'	(*Ctenomys conoveri*)
'wax palm, karanda'y'	(*Copernicia alba*)
'wolf-fish'	(*Hoplias malabaricus*)
'yerba mate'	(*Ilex paraguariensis*)

Appendix C

Elicitation Word-List

C.1 Words elicited from Toba-Enenlhet

The following elicitation prompts were used to elicit words in Angaité that were elicited in Toba-Enenlhet during my July 2018 pilot trip to Nueva León. Words were prompted in Guaraní or Spanish, with Toba-Enenlhet only used when speakers (or I) did not know a word in Guaraní and my consultants did not know the Spanish. Guaraní and Spanish words are written in the languages' practical orthographies.

Spanish	**Guaraní**	**Toba-Enenlhet**
fuego	tata	[tała]
leña	jepe'a	[jentapaʔ]
yo	che	[koʔo]
tú	nde	[łejap]
noche	pyhare	[pełasep]
día	ára	[aknem]
tarde	ka'aru	[jenoktezo]
gato	mbarakaja	[tenoq]
perro	jagua	[semheŋ]
mañana (morning)	pyhareve	[soseke?]
loro	pancho	[pełpo]
mañana (tomorrow)	ko'ẽrõ	[soseka?]
gallina	ryguasu	[tapeʔe]
lorito		posek
tereré	ka'a	[kaʔa]
barro	tuju	[jalpaʔ]
lejos	mombyry	[tajipeʔ]
humo		[eten]
polvo		[łepop aktek]
comida	tembi'u	[nentomaʔ]
carne	so'o	a͡petik
carbón		[mełma]
cenizas		[tahap]

gallo		[tapeʔe kenoʔ]
varón	kuimba'e	[apkenoʔ]
hembra	kuña	[kiluana]
vaca	vaka	[waka]
burro		[malekaʔ]
venado	guasu	[popijeʔ]
zorro		[maliŋ]
conejo	tapiti	[tomahaŋ]
	charata	[kenaten]
mono	ka'i	[jatsene]
víbora	mbói	[menasma]
araña	ñandu	[sewalaq]
mosquito	ñati'u	[paga]
lechuza		[sewaja]
rana		[naptektiŋ]
pez	pira	[kelasma]
tigre		[kemhawaʔ]
puma		[meiwaʔ]
carpincho	capybara	[itaptomahaŋ]
oso hormiguero		[anajim]
	tagua	[poma]
chancho	kure	[jitapoma]
	jakaré	[jitapijim]
tortuga	karumbe	[nataʔ]
armadillo	tatú	[ɲalwa]
flamenco		[kakɛŋ]
picaflor	majnomby	[sensaʔap]
largarto		[pehe]
	tatú bolita	[sepop]
paloma		[konemaqtek]
zorrillo		[mepa]
tapir		[pakɬama neteŋ]
avestruz	ñandu	[peɬapen]
buitre		[mamaʔ]
tokai		[popen]
mes	jasy	[pelten]
ahora	ko'ãga	[koɬawenak]
viento norte		[aɬkajam]
viento sur		[ijam]
sur		[nepijam]
amanecer	ko'ẽ	[atsaja]
no	nahániri	[haweʔ]

bueno/hermoso	porã	[tase]
mal	ñaña	[maŋkolatsema]
comienzo	pyrũ	[ampak^{h}aʔ]
palo santo		[memoŋ]
	capiatí	[meneke]
	yvyra hũ	[jajit]
quebracho colorado		[maseʔ]
poroto del monte		[tawa]
poroto	kumanda	[kalike]
sandia		[samania]
zapallo		[jeqtepa]
papa		[jampijeʔ]
batata	jety	[pijeʔ]
manioca	mandi'o	[sepo]
arroz		[pataktek]
algodón	mandyju	[jetepepa]
mani	manduvi	[makua]
maiz	avati	[alatse]
flor	yvoty	[iɬnemoq]
palmera	karanda'y	[ała]
palo blanco		[anawaʔ]
pallo borracho	samu'u	[anamoq]
pasto	kapi'i	[paʔat]
viento	tu	[ałkajam]
luna	jasy	[pelten]
semilla		[aktek]
estrella		[jalwa]
piedra	ita	[metajmoŋ]
camino	tape	[amaj]
sombra	kuarahy'a	[peskeska]
cielo		[neten]
laguna salada		[ajonkomeɬ]
nube		[jepopaja]
agua	y	[jemen]
montaña	yvyty	[metajmon]
monte		[naɬma]
campo	okara	[jokaɬma]
huevo	ryguasu rupi'a	[tapeʔe akpok]
pato	ype	[metektiŋ]
abeja		[jitajuhene]
vizcacha		[paimoŋ]
hombre	kuimba'e	[enenɬet]

mujer	kuña	[kiluana]
compañero	tapicha	[ałeɬma]
ciudad	tetã	[jawək^{h}atema]
pueblo	tenda	[negawaka, awak^{h}a]
hamaca	kyha	[tɛteko]
cama	tupa	[nentanema]
cuarto	koty	[nentijanmak^{h}aʔ]
casa	hóga	[nenemak^{h}aʔ]
puerta	okẽ	[atoŋ]
techo		[nenpakɬanma]
pared		[neŋhaɬtama]
ropa	ao	[astaɬnema]
camisa	sai	[nentaɬlema]
zapato	sapatu	[nenaktetomankok]
tela		[nepoʔak]
canoa		[lomajen]
auto	auto	[alto]
plata	viru	[patakom]
sueldo		[nenɬewaja]
idioma	ñe'ẽ	[nempajowoma]
trabajo	mba'apo	[nenmakmekama]
mosca	mberu	[japa]
hormiga		[hapoŋ]
pantalla (fire starting fan)		[ɬempenɪk]
ratón	anguja	[wapa]
planta	yvyra	[niŋanma]
río	ysyry	[sɪŋameʔ]
mar	y guasu	[alwata]
miedo	kyhyje (v)	[niŋajʔa]
madre	sy	[eŋken]
padre	tuva	[tata]
hermana		[ijaɬen]
hijo/a		[iʔitka]
niño/a	mitã	[nematka]
tío		[tatapiaɬen]
tía		[memeŋaten]
abuelo	taita	[sɛtaʔ]
abuela	jaryi	[sɛmaʔ]
nieto		[ɛtawɛn]
esposa		[ɪtawa]
famila	ógaygua	[enketkak]
hogar	óga	[nenemʌkʌt]

bebé (masc.)	mitã'i	[nɛmatkakɛtkok]

C.2 Intercontinental Dictionary Series: Lengua

The following elicitation prompts were used to elicit words in Angaité based on the Key and Comrie (2015) list for Lengua (Enlhet?). Words were propmted in Enlhet. If speakers did not recognize the word, they were additionally prompted with Guaraní or Spanish. Translations from English to Spanish or Guaraní were done ad hoc during elicitation. Words in Key and Comrie (2015) are written in practical orthography, which is used here. This orthography is the same as the Angaité orthography used in this paper except that <w> represents [w] instead of <v>.

English	**Enlhet**	**English**	**Enlhet**
sand	yinkamith	mountain, hill	inkyilhe
lake	ikyakyeng waiam	river	watsam, elwata
swamp	sangye	woods, forest	nathma
sky	netin	star	apyowa
thunder	takha	rainbow	pimhit
dew	yalyi	fog	tiyametin
ashes	tahap	embers	atith
firewood	yantapuk	boy	wukmaak
girl	inkilana etkuk	ancestors	yehengkyaa
orphan	hkuk	widow(er)	lhintampe
goat	yatai	bird	nata
heron	pana	bat	mepop
rabbit	tomahang	fish scale	impehik
insect	askuk	ant	topoli
beeswax	papa	grasshopper	sowa
bodo	yoklhoho	blood	ema
ear	-haikuk	eye	-aktik
nose	-waihik	tooth	-maak
neck	-yispuk	arm	-ektong
hand	-mik	foot	-minik
knee	-tapnik	heart	wanneya
die/dead	kyitsipkyi	medicine	panakte
bowl	hopuku	knife	sowu
meal	ninto	grease/fat	pithmuk
thread	tama	necklace	momatik
door	atong	tobacco	hena

palm tree	alha	gourd	kyaiyi
fan	alha awa	canoe	namuk
before, front	mamyi		
large, big	yowea	one	lhama
two	anit	three	antan lhama
year	apyiam	sweet	atsik
salty	yasik	hot	atehe
name	wisaia		
stranger	pok enlhit	spear	hewa

C.3 Intercontinental Dictionary Series: Sanapaná

The following elicitation prompts were used to elicit words in Angaité based on the Key and Comrie (2015) list for Sanapaná. Prompts were given in Sanapaná with additional prompting in Guaraní or Spanish if speakers did not recognize a word. Translations from English into Spanish and Guaraní were done ad hoc during elicitation. The IDS list for Sanapaná is done in the IPA, though I think that <y> represents [j] rather than [y].

English	**Sanapaná**	**English**	**Sanapaná**
bamboo	[yaamaamok]	sugar cane	[yampaʔat walay yawhan]
knife	[sowow]	nail	[laawa]
net bag	[soŋseŋheʔ]	swim	[neŋmaam yeŋmen]
walk	[neŋɬeŋaam naaɬpop]	come	[neŋɬeŋayaamʔ]
enter	[nentaɬnama]	afternoon	[taɬnaamʔ]
hour, clock	[textex]	white	[pak]
black	[akpaasyam]	red	[akyeɬwaseem]
yellow	[tap, akyaatektama]	good	[sas, aktaʔmalma]
bad	[neŋmasom]	difficult	[yemteɬeʔ]
easy	[kamanyeheme]	flute	[yaamaamok]
guest	[maykaʔa]	god	[eŋyaapam]
land	[amjep]	cave	[maaɬek]
lake	[alaʔɬok]	lake	[saaŋa]
lake	[paŋkaɬwa]	river	[waʔatsam]
spring	[jaamelket]	stone	[apjawʔa]
shade	[peskeska]	air	[ajennamaam]
cloud	[jephopaj]	fog	[teejam eeten]
rain (n)	[akmaamaj]	flame	[alajwoom]
burn (vtr)	[neŋwatnama]	charcoal	[aateɬ]
boy	[jooŋes]	husband	[entaawaʔ]
mother	[aŋken]	son, daughter	[eŋketka]

offspring	[entawaan]
youner sister, brother	[eŋjaaɬeŋ]
grandfather	[eŋjaataʔ]
pasture	[waetke ɬaŋkok]
pig	[paawa]
owl	[sawaaja]
mouse, rat	[wesetse]
lion	[neptaana]
insect	[askok]
centipede	[seŋjawan]
ant	[haapoŋ]
wasp	[mahaʔ]
fly (n)	[yaapa]
mosquito	[paaja]
snake	[nemesma]
snail	[akpoŋ]
lizard	[peehe]
lizard	[jaatepooma]
hair	[eŋwaʔ]
tail	[ajpakjek]
head	[empaʔtek]
eyebrow	[entekwahan]
elbow	[empeŋkeptek]
claw	[apheok]
udder	[neme]
pot	[wajka]
fork	[eŋepheok]
meat	[aapetek]
fruit	[akjeɬna]
chili pepper	[naatekhet]
fermented drink	[anmen]
clothes	[apaawa]
fur	[ajempehek]
thread	[taamaʔ]
skirt	[neŋjapmaam]
hat	[aaɬaʔwaʔ]
house	[teŋma]
hut	[haameɬ]
wall	[teŋma pwahak]
roof	[teŋma neeten]
garden	[amjep]
hoe	[aɬpehek]
older sister/brother	[eŋjaaɬoʔ]
old man	[kelaphopa]
grandmother	[eŋjaamaʔ]
cow	[waetke]
goose	[jaamtataʔa]
cat	[keɬkaaɬek]
mouse, rat	[pomoʔo]
deer	[lanaep]
louse	[peetem]
spider	[sewaalak]
spider	[melanma]
wasp	[kenaawet]
fly (n)	[semsek]
worm	[melʔaok]
butterfly	[seleklek]
frog	[semheʔ]
lizard	[pooma]
turtle	[jaatanaʔtaʔ]
horn	[aŋkeepeetek]
back	[eŋkapok]
face	[neŋaaʔat]
tongue	[aŋaɬkok]
finger	[eŋpehek]
feather	[aawaʔ]
hunger	[maek]
plate, bowl, dish	[poko]
sausage	[ajeŋmenek]
bean	[kelajke]
nut	[makwa]
butter	[peɬmok]
egg	[appok]
wool	[nepkeeseʔ pwaʔ]
needle	[ketaamaʔ]
weave	[neŋootma]
shoe	[tejajpehek]
tattoo	[nentaaɬamaawoʔ]
house	[neŋma]
floor	[naaɬpop]
chair	[nentahaneem]
post, pole	[meemoŋ]
spade	[menmeʔ]
rake	[eŋepheok]

plant	[mojanma]	plant root	[awhak]
root	[ajkepmenek]	branch	[aɬkapok]
leaf	[aawaʔ]	flower	[eɬnamok]
tree stump	[mompheje?]	sap	[anjeʔ]
palm tree	[paʔaŋ]	citrus	[sooljaatektama]
fish poison	[seŋwaj]	mushroom	[hapenenjeʔ]
rope, cord	[taamaʔ]	rope, cord	[aaɬaataʔaok]
knot	[asomnek]	[saw] (n)	[ajpakjek]
clay	[hajmakok]	fan	[neŋmakhemmaɬkaʔ]
thing	[aʔsok]	money	[soljaʔjem]
wages	[nenɬoj]	wages	[neŋmaam]
behind	[taap]	slow	[nentaʔmelseejkamʔol]
end	[neeɬaʔ]	soon	[keɬwooʔ]
night	[aɬtaʔa]	year, summer	[mokhetma]
sweet	[akmaases]	salty, bitter, sour	[akmaska]
sharp	[was]	hot	[koh]
beautiful, good	[sas]	bad	[neŋmasom]
difficult	[jemteɬeʔ]	guest	[majkaʔa]
host	[neŋaokɬaʔ]	bow	[peɬpaaʔaɬ]
arrow	[jaʔaŋka]		

C.4 Chaco-specific flora and fauna

The following words from a list of Chaco-specific flora and fauna which I created during my pilot trip in summer 2018 were used to elicit responses in Angaité. The words were prompted with the Guaraní name of the animal, followed by a description or picture (if available) if speakers did not recognize the animal's name in Guaraní. Where known, this reference list includes the term in Spanish.

Guaraní	**Spanish**	**Guaraní**	**Spanish**
capybara	carpincho		tucatuca
	puma	jaguané	zorrillo
mboreví	tapir	taguá	chancho quimilero/pecarí del Chaco
aguará guazu	zorro crinudo	aguara'i	zorro gris/zorro de monte
tañicati	pecarí labiado	guazu vira	venado gris
guazy pytã	venado rojo	karayá	mono aullador
ka'i pyhare	mono de la noche	jakaré hũ	
ñeromi		teyu pytã	
jarará		kururú chini	escuezo chaqueño
kururú pytã		ynambú	

yabiru	tuyuyú cuartelero	charata
ñacurutú	buho	urundey
	palo santo	karaguatá
mburucujá	(passion flower)	karandá

References

Amarilla, Deisy. 2008. *Agtemágháhta apqueh hleanma: Pueblo Sanapaná*. Asunción: Biblioteca Paraguaya de Antropología.

Amarilla, Juan Pablo. 2006. *Pueblo Angaité: Memoria Histórica*. Asunción.

Arenas, Pastor. 1981. *Etnobotanica Lengua-Maskoy*. Buenos Aires: Fundación para la educación, la ciencia y la cultura.

Bethin, Christina Y. 2011. Polish syllable structure. In: *Blackwell Companion to Phonology*. Ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice. Vol. 5. Hoboken, NJ: John Wiley & Sons.

Breen, Gaban and Rob Pensalfini. 1999. Arrernte: A langauge with no syllable onsets. *Linguistic Inquiry* 1, pp. 1–25.

Campbell, Lyle. 2012. Typological Characteristics of South American Indigenous Languages. In: *Indigenous Languages of South America: A Comprehensive Guide*. Ed. by Lyle Campbell and Verónica Grondona. Berlin: De Gruyter, pp. 259–330.

Campbell, Lyle. 2013. Language Contact and Linguistic Change in the Chaco. *Revista Brasileira de Linguística Antropológica* 5.2, pp. 259–291.

Campbell, Lyle and Verónica Grondona. 2010. Who speaks what to whom? Multiculturalism and language choice in Misión La Paz. *Language in Society* 39, pp. 617–646.

Carlisle, Robert S. 2001. Syllable structure universals and second language acquisition. *International journal of English studies* 1 (1), pp. 1–19.

Davidson, Lisa and Daniel Erker. 2014. Hiatus resolution in American English: the case against glide insertion. *Language* 90 (2), pp. 482–514.

Domaniczky, Mariana Franco and Gladys Imaz. 2006. *Angaite-Koahlvok-Las Voces de un Pueblo*. Asunción: CEADUC.

Durante, Santiago. 2011. Las lenguas del Gran Chaco: Situación socio-lingüística y políticas lingüísticas. *Language Design* 13, pp. 115–142.

Fabre, Alain. 2018. *Los enlhet-enenlhet del Chaco Paraguayo*. Available at http://www.ling.fi/Entradas%20diccionario/Dic=EnlhetEnenlhet.pdf (2018/11/10).

Glauser, Marcos. 2019. *Angaité's Responses to Deforestation: Political ecology of the livelihood and land use strategies of an indigenous community from the Paraguayan Chaco*. Zürich: Curupira.

Gobierno Nacional. 2012. *Atlas de Comunidades de Pueblos Indígenas en Paraguay 2012: Familia lingüística Lengua Maskoy*. Asunción: Dirección General de Estadística, Encuestas y Censos.

Gomes, Antonio Silva. 2009. Sanapaná: Um povo indígena paraguaioi. *Sintesis*, pp. 360–385.

Gomes, Antonio Silva. 2013. Sanapaná uma lingua maskoy: Aspectos gramaticais. PhD thesis. São Paulo: Universidade estadual de campinas, Instituto de estudios da linguagem.

González, Hebe A. 2014. Procesos fonológicos como rasgos areales: el caso de la palatalización en las lenguas chaqueñas. *LIAMES* 14, pp. 11–39.

González, Hebe A. 2015. El Chaco como área lingüística: una evaluación de los rasgos fonológicos. In: *Language Contact and Documentation*. Ed. by Bernard Comrie and Lucía Golluscio. Berlin: De Gruyter Mouton, pp. 165–203.

Grubb, W. Barbrooke. 1911. *An unknown people in an unknown land: An account of the life and customs of the Lengua indians of the Paraguayan Chaco, with adventures and experiences during twenty years' pioneering and exploration amongst them*. London: Seeley & Co. Limited.

Hammarström, Harald, Robert Forkel, and Martin Haspelmath. 2019. *Glottolog 4.1*. Max Planck Institute for the Science of Human History. Accessed: 2019-10-09. https://glottolog.org.

Hu, Fang and Chunyu Ge. 2016. Vowels and diphthongs in Cangnan Southern Min Chinese dialect. In: *Interspeech* (San Francisco, CA), pp. 978–982.

Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23, pp. 225–257.

Hyman, Larry. 2010. Does Gokana really have no syllables? (Or: what's so great about being universal?) *UC Berkeley PhonLab Annual Report* 6.6.

Jaggers, Zachary Scott. 2018. Evidence and characterization of a glide-vowel distinction in American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9 (1), pp. 1–27.

Kalisch, Hannes. 2009. Los constituyentes de la cláusula enlhet (enlhet-enenlhet): Esbozo de una cláusula omnipredicativa. *Amerindia* 33/34, pp. 109–148.

Kalisch, Hannes. 2019. Orden de los argumentos en Enlhet. In: *La interfaz syntaxis-pragmática: Estudios teóricos, descriptivos y experimentales*. Ed. by Valeria A. Belloro. Berlin: De Gruyter, pp. 143–177.

Key, Mary Ritchie and Bernard Comrie, eds. 2015. *IDS*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://ids.clld.org/.

Levi, Susannah V. 2011. Glides. In: *Blackwell Companion to Phonology*. Ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice. Vol. 1. Hoboken, NJ: John Wiley & Sons.

Nobbs-Thiessen, Ben. 2017. Reshaping the Chaco: Migrant Foodways, Place-making, and the Chaco War. *Journal of Latin American Studies* 50, pp. 579–611.

Powys, E.L.G. 1929. A Practical Lengua Grammar. MA thesis. Oxford: Oxford University.

Rivera Castillo, Yolanda. 2017. Hiatus resolution in Puerto Rican Spanish and language typology. In: *Current research in Puerto Rican linguistics*. London: Routledge, pp. 23–42.

Sánchez Miret, Fernando. 1998. Some reflections on the notion of diphthong. *Papers and studies in contrastive lingusitics* 34, pp. 27–51.

Susnik, Branislava. 1958. Eenlhet Appaiwa. Lengua-Maskoy: estructura gramatical. *Boletín de la Sociedad científica del Paraguay 2*, pp. 1–92.

Susnik, Branislava. 1977. *Lengua-maskoy, su hablar, su pensar, su vivencia*. Asunción: Museo Etnográfico Andrés Barbero.

Unruh, Ernesto and Hannes Kalisch. 1997. *Moya'ansaeclha' Nengelpayvaam Nengeltomha Enlhet*. Yalve Saanga: Colección Biblioteca Paraguaya de Antropología Vol XXVII (CEADUC).

Unruh, Ernesto and Hannes Kalisch. 1999a. Defender nuestro idioma. El idioma en la redefinición de la identidad indígena: El caso de los Enlhet. *Acción: revista paraguaya de reflexión y diálogo* 181, pp. 34–36.

Unruh, Ernesto and Hannes Kalisch. 1999b. Escribir en idioma enlhet para hacer crecer lo propio. *Suplemento antropológico* 34.1, pp. 101–176.

Unruh, Ernesto and Hannes Kalisch. 2003. Enlhet-Enenlhet: Una familia lingüística chaqueña. *Thule, Rivista italiana di studi americanistici* 14/15, pp. 207–231.

Unruh, Ernesto, Hannes Kalisch, and Manolo Romero. 2003. *Enenlhet Apaivoma. Nentengiai'a nengiangveiakmoho neliatekamaha Enenlhet apaivoma: Guía para el aprendizaje*

del idioma materno toba. Ya'alve-Saanga/Chaco paraguayo: Nengvaamequescama Nempayvaam Enlhet. www.enlhet.org.

van Gysel, Jens. 2017. Temporal Predicative Particles in Sanapaná and the Enlhet-Enenlhet Language Family (Paraguay): A Descriptive and Comparative Study. MA thesis. Leiden: Leiden University Centre for Linguistics.

Villagra Carrón, Rodrigo. 2014. *Meike makha valayo: No habían paraguayos*. Vol. 98. Asunción: Biblioteca Paraguaya de Antropología.

Wheeler, Paige Erin. 2019–. *The Angaité Collection of Paige Erin Wheeler*. https://ailla.utexas.org. The Archive of the Indigenous Languages of Latin America. Access: public. PID ailla:267815.

VITA

Paige Erin Wheeler was born in Sturgis, Michigan and grew up near Bristol, Indiana. She first visited Paraguay in 2011 on a high school trip. She began attending Grinnell College in 2012, receiving a B.A. in Spanish, with honors, and a concentration in Linguistics in 2016. After completing her B.A. she worked for a year at the Northwest Immigrant Rights Project in Seattle, WA, where she coordinated intake interviews for the children and youth services unit. She entered a PhD program in Linguistics at the University of Texas at Austin in 2017 and has been conducting fieldwork in Paraguay on Enlhet-Enenlhet languages since 2018.

Address: pewheeler@utexas.edu

This report was typeset with LaTeX by the author.