Phonological Reconstruction of Proto-Kampa Consonants*

Department of Linguistics, University of California, Berkeley

Eric Chen

June 2019

1 Introduction

1.1 Goals and Outline

The primary goal of this article is to apply the comparative method to the reconstruction of Proto-Kampa consonants. Specifically, this entails creating cognate sets using the six Kampa varieties (Nomatsigenga, Ashéninka, Pajonal, Asháninka, Kakinte, and Matsigenka) in my data and generating correspondence sets for the consonants. These correspondence sets form the basis for my reconstruction of the Proto-Kampa consonant phonemic inventory and the sound changes that resulted in the diversification of consonant phonemes in the daughter varieties.

A secondary goal of this article is to demonstrate the efficacy of newly developed tools for computational historical linguistics, namely, the LingPy and LingRex Python libraries and the Edictor program, all written by Johann-Mattis List. These technologies greatly enhanced the efficiency of my work without sacrificing accuracy. It is clear that the field of historical linguistics stands much to benefit from the increased integration of such computational tools.

In the remainder of Section 1, I provide a brief overview of relevant background information about the Kampa family. I then lay out the phonemic inventories of the six varieties in my data set and discuss previous work on Proto-Kampa reconstruction. Section 2 discusses the methodology of my research and describes how I used computational tools to help generate cognate sets and correspondence sets. In Section 3, I present my reconstructed Proto-Kampa consonant phonemic inventory and justify my decisions using evidence from the correspondence sets. Section 4 walks through the diversification of Proto-Kampa consonants in the daughter varieties through various sound changes. It also contains cognate set data demonstrating the sound changes and discusses irregularities and exceptions in the data. Finally, Section 5 summarizes my research and suggests possible avenues for future study.

1.2 The Kampa Family

The Kampa family is a branch of the Arawak family, which contains 40 living languages spoken across northern South America and the Caribbean (Aikhenvald 1999, p. 65). Most scholars place the Kampa languages within the Southern Arawak grouping, which essentially comprises the Arawak languages

*This article is adapted from my undergraduate honors thesis completed in Spring 2019. It has been reformatted but no substantive changes have been made to its content. I would like to thank my advisor, Professor Lev Michael, for guiding me through the research and writing process and Zachary O'Hagan for providing insightful feedback on my colloquium presentation. I would also like to thank Johann-Mattis List for developing a free suite of tools for computational historical linguistics which I used in my research. I use the following abbreviations: PK = Proto-Kampa, Nom = Nomatsigenga, Ashé = Ashéninka, Paj = Pajonal, Ashá = Asháninka, Kak = Kakinte, Mat = Matsigenka.

spoken south of the Amazon River. Within Southern Arawak, Kampa has been grouped with the Purús and Bolivia-Parana families (Michael 2008, p. 219).

The Kampa languages are spoken throughout central Peru, from the eastern Andean foothills to the western fringe of the Amazon basin (Mihas 2017, p. 784). Scholars disagree on the identification and naming of Kampa varieties as well as their classification into languages and dialects (Michael 2008, p. 213) (Mihas 2017, p. 782). According to Michael, the commonly recognized "major varieties" are Nomatsigenga, Ashéninka, Pajonal Ashéninka (which I will henceforth refer to as Pajonal), Asháninka, Kakinte, Matsigenka, and Nanti (Michael 2008, p. 213). My reconstruction uses wordlists from all of these varieties except Nanti, which is closely related to Matsigenka (Michael 2008, p. 218). In the remainder of this article, I will use the term "variety" to avoid making any claims about languages or dialects. I also adopt the spelling conventions used by Michael for the names of the varieties.

It is clear that Ashéninka and Pajonal are closely related, and I will sometimes refer to Pajonal as a variety of Ashéninka. A priori, I make no other assumptions or claims about the relationships among the varieties.

The Kampa population includes 97,477 ethnic Asháninka and Ashéninka, 11,279 Matsigenka speakers, 8,016 Nomatsigenga speakers, 439 people identifying as Kakinte, and no more than 450 Nantis (Mihas 2017, p. 784) (Michael 2008, p. 20). A number of Kampa varieties have very few speakers and are highly endangered (Mihas 2017, p. 784).

Kampa varieties have open word classes of verbs and nouns, closed classes of pronouns and adverbs, and adjectives, whose properties differ among Kampa varieties (Mihas 2017, p. 789).

Kampa varieties are highly synthetic and agglutinating (Mihas 2017, p. 788). A practical consequence of this fact is that verb roots and many noun roots cannot occur in isolation, so they are presented as inflected forms in the various dictionaries. The process of extracting the roots from these forms is non-trivial and is described in Section 2.1.

Kampa varieties have a syllable structure of (C)V(V)N (Mihas 2017, p. 787). N represents the placeless nasal phoneme, which can only occur in the coda and is the only consonant which can occur in the coda. There are no consonant clusters except for sequences of /N/ followed by a stop or affricate. See Section 1.3.1 for more information about the placeness nasal.

Stress is not contrastive in Kampa varieties. According to Mihas, "In nouns, stress is attracted to the penultimate syllable, in verbs it is based on disyllabic iambic feet, with the main stress assigned to the final foot" (Mihas 2017, p. 787). Stress is not marked in the data presented in this article.

1.3 Phonological Inventories

Here I delineate the phonological inventories of the Kampa varieties in my data set. These inventories are constructed from descriptions in my lexical sources (Shaver 1996, p. 13) (Payne 1980, pp. 7–10) (Heitzman and Doerksen 2017) (Kindberg 1980, p. 6) (Snell 2011, pp. 17–18) as well as other phonological descriptions of the varieties (Lawrence 2013, p. 8) (Payne 1981, p. 59) (Swift 1988, p. 100) (Dohn 2017, p. 7).

According to Mihas, the consonant phoneme inventories of Kampa varieties generally share the following similarities with those of other Arawakan languages: moderate degree of variation in the place of articulation, no phonemic voice distinctions in dental stops, palatalized dental stops, two sibilant fricatives, at least one rhotic, and the palatal nasal (Mihas 2017, p. 785). However, I do not reconstruct palatalized dental stops or the palatal nasal for Proto-Kampa (see Section 3).

Mihas implies that in all Kampa varieties, the vowels /a/, /e/, /i/, and /o/ have phonemic long counterparts (Mihas 2017, p. 786). Thus, I list them for all six varieties even when I do not have a source explicitly saying that they are phonemic in a particular variety. In all of the lexical data in Section 4, long vowels are written as a double vowel. The rationale for this is explained in Section 2.2.3.

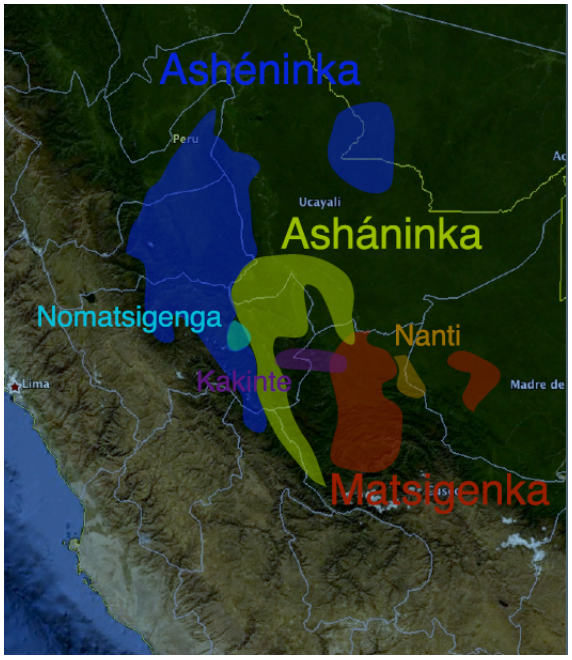

Figure 1: Map of Kampa varieties (Michael 2011)

1.3.1 The Placeless Nasal

Nomatsigenga, a variety of Ashéninka, Kakinte, and Nanti have been analyzed as having a nasal phoneme, symbolized /N/, which is unspecified for place of articulation (Lawrence 2013, p. 9) (Payne 1981, p. 62) (Swift 1988, p. 99) (Michael 2008, p. 223). This segment only occurs in syllable codas before a stop or affricate in the following onset, from which it acquires its place features.

Michael makes several arguments for positing /N/ as a phoneme in Nanti rather than analyzing all instances of it as /m/ or /n/. First, when it occurs at the end of a morpheme, there is no basis for choosing one place of articulation for it over another in the underlying form. Second, /N/ deletes before vowels, unlike /m/ or /n/. Finally, if we do not posit /N/, then [ŋk] sequences within a morpheme would have to be analyzed as /ŋk/, but there is no evidence for the phoneme /ŋ/ (Michael 2008, pp. 223–224). In her analysis of Nomatsigenga, Lawrence further argues why /Np/, /Nt/, and /Nk/ sequences should not be treated as /ᵐp/, /ⁿt/, and /ᵑk/ (Lawrence 2013, p. 10). These arguments for /N/ essentially apply to all Kampa varieties. Thus, I include it in all of the following inventories whether or not it is mentioned in any phonological description of the variety. However, in all of the lexical data presented in Section 4, /N/ is place-assimilated and written as /m/, /n/, /ɲ/, or /ŋ/.

1.3.2 Nomatsigenga Inventory

Lawrence gives the following inventory of Nomatsigenga consonant phonemes (Lawrence 2013, p. 8). This is in accordance with the description provided by my lexical source (Shaver 1996, p. 13). I defer to Lawrence's judgment in calling /n/, /t/, /ts/, and /s/ dental rather than alveolar and in grouping /ɾ/ with the (alveo)palatal segments rather than the dental segments. I do not have any justification for these decisions but they are of little bearing to the reconstruction.

	labial	dental	(alveo)palatal	velar	glottal	unspecified
nasal	m	n		ŋ		N
stop	p b	t		k g		
affricate		ts	tʃ			
fricative		s	ʃ		h	
liquid			ɾ			
glide			j			

Table 1: Nomatsigenga consonant phonemes

Lawrence gives the following inventory of Nomatsigenga vowel phonemes, but she writes the high central unrounded vowel as /ɨi/ (Lawrence 2013, p. 8). I have chosen to represent this vowel as /ɨ/ because the pronunciation guide of my lexical source does not indicate that it is a diphthong (Shaver 1996, p. 13). In any case, this vowel phoneme is monomoraic and does not have a long counterpart (Lawrence 2013, p. 12). /eː/ appears to be rare and it is not attested in my data.

	front	central	back
high	i iː	ɨ	
mid	e eː		o oː
low		a aː	

Table 2: Nomatsigenga vowel phonemes

The Nomatsigenga consonant inventory differs from the general Kampa profile in its inclusion of /ŋ/, while in other varieties [ŋ] only occurs as an allophone of /N/. Nomatsigenga also contrasts high tone and no tone, which is unique among Kampa varieties (Mihas 2017, p. 787). I represent high tone with an acute accent on the mora which bears it.

1.3.3 Ashéninka Inventory

The following Ashéninka consonant inventory is constructed from the pronunciation guide in my lexical source (Payne 1980, pp. 7–10). /N/ has been added for reasons described above and /β^j/ has been removed as it only occurs once in the dictionary in a PI-Nevati dialect form. See Section 4.1.4 for more information relating to /β^j/.

This inventory is similar to the one given by Payne for the Apurucayali variety of Ashéninka (Payne 1981, p. 59). That dialect does not have /m^j/, /p^j/, and /k^j/, and Payne analyzes the the equivalents of /β/ and /ɣ/ as glides instead of fricatives.

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m m^j	n	ɲ			N
stop	p p^j	t t^h		k k^j		
affricate		ts ts^h	tʃ $tʃ^h$			
fricative	β	s	ʃ	ɣ	h	
liquid		ɾ	$ɾ^j$			
glide			j			

Table 3: Ashéninka consonant phonemes

The following vowel inventory is shared by Ashéninka, Pajonal, Asháninka, and Kakinte. These varieties have merged the vowel corresponding to Nomatsigenga /ɨ/ and Matsigenka /ɯi/ with short /i/, leaving a system with four vowel qualities in short-long pairs.

	front	central	back
high	i iː		
mid	e eː		o oː
low		a aː	

Table 4: Ashéninka, Pajonal, Asháninka, and Kakinte vowel phonemes

The phonology of Ashéninka diverges from the general Kampa profile in that it has contrastive aspiration and a large set of palatalized consonants.

1.3.4 Pajonal Inventory

The following Pajonal consonant phonemes are constructed from a description of the Pajonal alphabet, with the addition of /N/ (Heitzman and Doerksen 2017). This document includes /hʲ/ (written <jy>), but <jy> is not attested in the dictionary so I do not include it here.

Following Payne's analysis of the related variety of Apurucayali Ashéninka, I place /tʲ/ and /ɾʲ/ in the (alveo)palatal column. This inventory differs from that of Ashéninka in that it contains /tʲ/ and /tʃ/ instead of /tʃ/ and /tʃʰ/, /w/, and /wʲ/ instead of /β/, and lacks /s/. The presence of the palatalized /wʲ/ is significant, but the choice of /w/ rather than /β/ is an arbitrary one on my part, as [β] and [w] are allophones of this phoneme in both varieties (Payne 1980, p. 10) (Heitzman and Doerksen 2017, pp. 4–5).

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m mʲ	n	ɲ			N
stop	p pʲ	t tʰ	tʲ	k kʲ		
affricate		ts tsʰ	tʃ			
fricative			ʃ	ɣ	h	
liquid		ɾ	ɾʲ			
glide	w wʲ		j			

Table 5: Pajonal consonant phonemes

1.3.5 Asháninka Inventory

The following consonant phoneme inventory of Asháninka is constructed from the description of the orthography in Kindberg's dictionary (Kindberg 1980, p. 6), with the addition of /N/.

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m	n				N
stop	p	t		k		
affricate		ts	tʃ			
fricative	β	s	ʃ		h	
liquid		ɾ				
glide			j			

Table 6: Asháninka consonant phonemes

Asháninka is unique among the varieties in my sample in that it lacks a voiced velar segment.

1.3.6 Kakinte Inventory

Swift gives the following consonant inventory for Kakinte (Swift 1988, p. 100). In personal communication, O'Hagan noted that Swfit's use of /ɣ/ instead of /g/ is controversial and not solidly grounded

in phonetic facts, as this phoneme does not actually surface as [ɣ] in any environment.

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m	n	ɲ			N
stop	p	t	tʲ	k		
affricate		ts	tʃ			
fricative	β	s	ʃ	ɣ	h	
liquid		ɾ	ɾʲ			
glide			j			

Table 7: Kakinte consonant phonemes

Kakinte is like Ashéninka and Pajonal in that it has a large number of (alveo)palatal segments, but unlike those varieties it does not have any palatalized labial or velar consonants.

1.3.7 Matsigenka Inventory

The following Matsigenka consonant phonemes are largely constructed from the description in Snell's dictionary, with the addition of /N/ (Snell 2011, pp. 17–18). I also add /ɾʲ/, which is included in Dohn's phonological sketch but does not appear in my data (Dohn 2017, p. 7). I learned from personal communication with Michael that this is a very rare phoneme. I diverge from Dohn in grouping /tʲ/ and /ɾʲ/ in the (alveo)palatal rather than alveolar place of articulation, only because this is more consistent with descriptions of other Kampa varieties (Payne 1981, p. 59) (Swift 1988, p. 100) (Michael 2008, p. 221). This decision does not affect my reconstruction.

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m	n	ɲ			N
stop	p b	t	tʲ	k kʲ g		
affricate		ts	tʃ			
fricative		s	ʃ		h	
liquid		ɾ	ɾʲ			
glide			j			

Table 8: Matsigenka consonant phonemes

Like Nomatsigenga, Matsigenka has a monomoraic high non-front vowel (Mihas 2017, p. 786). I represent this segment as /ɯi/, following the phonetic description in Snell's dictionary and Michael's description of the similar variety Nanti (Snell 2011, p. 18) (Michael 2008, p. 225). /eː/ and /oː/ appear to be rare and are not attested in my data.

	front	central	back
high	i iː		ɯi
mid	e eː		o oː
low		a aː	

Table 9: Matsigenka vowel phonemes

Like Ashéninka and Pajonal, Matsigenka has a palatalized velar consonant and many (alveo)palatal consonants, but it lacks palatalized labial consonants.

1.4 Previous Work on Proto-Kampa Reconstruction

Here I describe previous work on the reconstruction of Proto-Kampa phonology and morphology. There has been no comprehensive reconstruction of Proto-Kampa phonology, and as of 2008 "there has been no application of the comparative method to the varieties of the Kampan family" (Michael 2008, p. 213).

1.4.1 Heitzman's Phonological Reconstruction

In 1973, Heitzman reconstructed the alveolar and alveopalatal stop and fricative phonemes in Proto-Kampa as **t*, **t^h*, **ts*, and **s* (Heitzman 1973, p. 12). This reconstruction is based on a comparison of wordlists from the following Kampa varieties: Campa along the Apurimac River, Campa of Baños, Campa of Cahuapanas, Campa of Chicosa, Campa of Cocari, Campa along the Ene and Tambo Rivers, Campa of Alto Irruya, Campa of Nivati, Nomatsiguenga Campa, Campa of Obenteni, Campa of Puerto Davis, Campa of Shinipo, Campa of Satipo, Campa of Shiringamaso, Campa of Tsireentsishavo, and Machiguenga (Heitzman 1973, p. 1). The precise mapping between these variety names and names used by later authors is unclear, but it appears that this sample includes numerous varieties of Ashéninka and Asháninka and lacks Kakinte and Nanti.

Heitzman's reconstruction differs significantly from mine, as I reconstruct **t*, **ts*, **tʃ*, **s*, and **ʃ* as the corresponding Proto-Kampa phonemes. Heitzman's reconstruction is not based on the comparative method, and she does not argue for why her reconstructed inventory is more likely than other possible inventories. Also, the wordlists that she relies on are of questionable accuracy. For example, they include the sequence /t^{h}i/ in varieties of Ashéninka for words where Payne's later dictionary uses both <tzi>, representing /tsi/ (compare /koβit^{h}i/ (Heitzman 1973, p. 53) and <covitzi> (Payne 1980, p. 54) meaning "POT") and <tsi>, representing /ts^{h}i/ (compare /aβot^{h}i/ (Heitzman 1973, p. 53) and <avotsi> (Payne 1980, p. 33) meaning "PATH"). For these reasons, in Section 3 I will simply argue for the correctness of my reconstruction without attempting to explain why Heitzman arrived at a different result.

1.4.2 Michael's Phonological Reconstruction

In 2010, based on wordlists for Asháninka, Ashéninka, Kakinte, Matsigenka, Nanti, and Nomatsigenga, Michael et al. outlined the following sound changes in the Kampa branch (Michael et al. 2010):

1. Northern Kampan high vowel merger: **ɯi* > *i* in Ashéninka, Asháninka, and Kakinte.
2. Velar stop lenition: In Ashéninka, **g* is lost following *a* or between identical vowels; otherwise it becomes /j/. In Asháninka **g* becomes /j/ before *e*; otherwise it is lost.
3. Development of contrastive aspiration in Ashéninka: First, **ts* underwent a three-way split, becoming /t^h/ before *a* and *o*, /ts^h/ before *i*, and /tʃ/ before *e*. Then, **t* became /ts/ before *i*.
4. **ŋk* > *ŋ* and **mp* > *m* in Nomatsigenga, creating the phoneme /ŋ/.
5. Rounding assimilation and dissimilation in Nomatsigenga, the details of which are irrelevant to the reconstruction of consonants.

Based on these sound changes, Michael et al. reconstruct the following Proto-Kampa phoneme inventory (Michael et al. 2010, p. 8).

	labial	alveolar	alveopalatal	velar	glottal
nasal	m	n			
stop	p b	t		k g	
affricate		ts	tʃ		
fricative		s	ʃ		h
liquid		ɾ			

Table 10: Michael et al.'s Proto-Kampa consonant inventory

	front	central	back
high	i j	ɯi	
mid	e		o
low		a	

Table 11: Michael et al.'s Proto-Kampa vowel and glide inventory

In a subsequent presentation, Michael noted two other sound changes: **s* palatalizes to /ʃ/ before *i* in each Kampa variety except Nomatsigenga, and **s* additionally palatalizes to /ʃ/ before *e* in Ashéninka, Asháninka, and Kakinte. Also, Michael characterizes the commonality between the lenition of **g* in Ashéninka and Asháninka as **g* > ∅ / *V_V*, i.e. intervocalic deletion. He then uses shared innovations to posit the following internal structure of the Kampa family (Michael 2011). **ii* represents the same PK phoneme as **ɯi*.

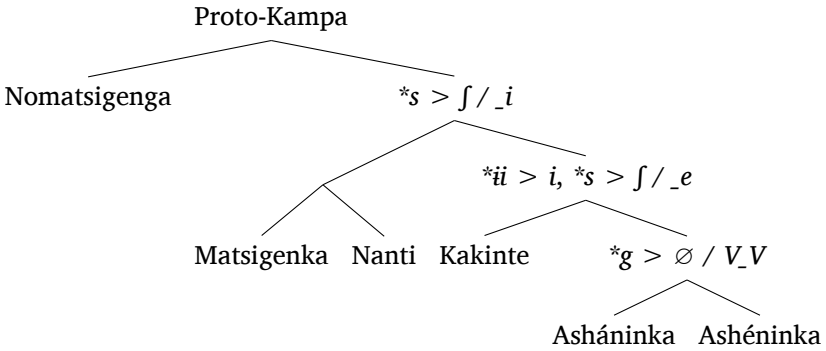

Figure 2: Internal structure of the Kampa family, option 1

This structure differs from the following structure earlier proposed by Michael, in which Nomatsigenga is grouped with Matsigenka and Nanti (Michael 2008, p. 218).

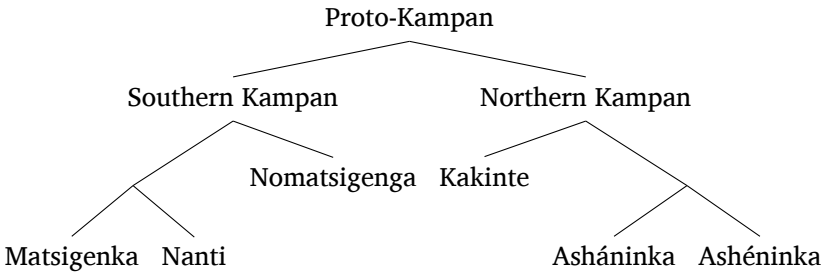

Figure 3: Internal structure of the Kampa family, option 2

1.4.3 Lawrence's Morphological Reconstruction

Using data from Nomatsigenga, Nanti, Matsigenka, Kakinte, Asháninka, and Ashéninka, Lawrence reconstructed varous aspects of Proto-Kampa verbal morphology including subject pronouns, object markers, number markers, directional markers, reality status markers, and valence-changing suffixes (Lawrence 2012). Lawrence notes that in addition to the sound changes identified by Michael, the synchronic and diachronic process of /h/-deletion is important for reconstruction, but she does not describe this process in detail (Lawrence 2012, p. 259). With regards to subgrouping, Lawrence points out that the distribution of the third-person non-masculine subject marker lends slight support to the first tree shown above, in which Nomatsigenga forms its own branch (Lawrence 2012, p. 269).

2 Methodology

In this section, I give a procedural overview of my research. The subsections on harvesting and processing wordlists describe processes that largely occurred in parallel for each Kampa variety in my sample, while the subsections on cognate set construction, correspondence set construction, and reconstruction apply to the combined data from all six varieties. Necessarily, the presentation in this section is more linear than the actual working process was. Also, for conciseness I omit some details and steps that have no bearing on the final results.

2.1 Harvesting Wordlists

My raw lexical data takes the form of a wordlist (perhaps better termed a list of roots) for each Kampa variety in my sample, built over a set of over 800 concepts. This set of concepts was developed by my advisor Lev Michael and contains general Swadesh List items as well as region-specific concepts like flora and fauna. The harvested wordlists contain various fields for each word, including but not limited to the orthographic form of the root, the Spanish translation given by the dictionary, the unified English translation (which I refer to as a concept), the grammatical category, the inflected form given by the dictionary if it differs from the root, and notes on borrowing, allomorphy, and dialectal variation.

For Nomatsigenga, Ashéninka, and Asháninka, I harvested the wordlists manually from the dictionaries: (Shaver 1996), (Payne 1980), and (Kindberg 1980) respectively. For Pajonal, I received an unpublished early version of a dictionary from Mary Ruth Wise through my advisor (Heitzman and Doerksen 2018). This dictionary was in the MDF file format (see (Coward and Grimes 2003) for more information about MDF), which allowed me to write a parser in Python to extract the fields automatically.

For Kakinte, my wordlist comprises data from fieldwork notes collected by Zachary O'Hagan (O'Hagan 2016). For Matsigenka, my wordlist comprises data from Snell's dictionary, harvested by Michael (Snell 2011).

In Sections 2.1.1 and 2.1.2, I describe the process of extracting noun and verb roots from inflected forms and issues in determining the root. This process applies to all varieties in my sample except Kakinte, whose data does not come from a dictionary.

2.1.1 Extracting Noun Roots

All Kampa varieties contain the three classes of inalienably possessed, alienably possessed, and non-possessable nouns (Mihas 2017, p. 790). The root of alienably possessed and non-possessable nouns can occur in its bare form, so this is the form given in the dictionaries. On the other hand, inalienably possessed nouns must occur with a possessive prefix or the suffix <-(n)tsi>, which signifies an unspecified possessor.

In cases where the dictionary provides a form with the unspecified possession suffix, I have stripped <-Ntsi> where possible and <-tsi> otherwise. This decision reflects the most common view among Kampanists, which is that <-Ntsi> and <-tsi> are indeed allomorphs of this suffix (Kindberg 1980, p. 18) (Mihas 2017, p. 790). However, Lawrence advocates for an alternative analysis for Nomatsigenga, claiming that nouns which seem to have the /-Ntsi/ allomorph actually have roots ending in /N/, which deletes when not followed by a suffix (Lawrence 2013, p. 11). For consistency, I have not adopted Lawrence's segmentation.

Cases in which the dictionary only supplies one or more forms with possessive prefixes were harder to deal with, because these prefixes often exhibit allomorphy which makes the root ambiguous (see (Shaver 1996, p. 34) for a list of allomorphs in Nomatsigenga). These cases were dealt with on a word-by-word basis, using other data from the same variety (e.g. occurrence of the root in compounds) as well as cognates in other varieties to inform the decision if possible. However, in some cases there is simply not enough information in the lexical source to make a conclusive judgment.

Another complication is that possessive prefixes trigger voicing of root-initial /p/ and /k/ in all Kampa varieties (e.g. /pagiɾo/ vs. /o-bagiɾo-te/ in Nanti and /ketsi-tsi/ vs. /na-getsi/ in Nomatsigenga (Michael 2008, p. 301) (Shaver 1996, p. 31)), although the specific values of the voiced counterparts differ among varieties. In cases where this alternation was displayed, the form with the voiceless consonant was chosen as it is the underlying form. However, if the only given form is a prefixed form with a root-initial reflex of PK **b* or **g*, the explicitly listed root was taken, but we have no way of knowing if it is the correct underlying form. This is the source of some voicing discrepancies in root-initial consonants in the data presented in Section 4.

2.1.2 Extracting Verb Roots

To preface the following discussion, I would like to note that all Kampa varieties epenthesize /a/ at a morpheme boundary between two consonants and /t/ at a morpheme boundary between two vowels (Lawrence 2013, p. 18) (Payne 1981, p. 111) (Swift 1988, p. 118) (Dohn 2017, p. 7).

For Ashéninka and Pajonal, extracting verb roots was relatively straightforward, as the dictionaries for those varieties list verbs in their nominalized form as root + <-aantsi>. Similarly, the Matsigenka dictionary lists verbs as root + <-agantsi>, with the caveat that root-initial vowels are deleted in this form, so they were inserted if they were found in other inflected forms. For these varieties, the only ambiguity occurs when the nominalized form ends in <-taantsi> or <-tagantsi>. In these cases, the dictionaries typically do not provide enough information to tell if the /t/ is epenthetic or part of the root. Generally, I treated it as epenthetic.

The dictionaries for Nomatsigenga and Asháninka do not give the nominalized form of a verb, but rather choose a seemingly arbitrary inflected form. Thus, in addition to the ambiguity regarding root-final /t/, this introduces other ambiguities regarding the boundary between the root and subject prefixes (which exhibit considerable allomorphy like possessive prefixes) and the boundary between the root and suffixes. In most cases, the authors of these dictionaries do provide a root, but their analysis is of questionable consistency and accuracy. Unclear roots for these varieties were handled on a case-by-case basis drawing on the expertise of my advisor Lev Michael. Instances in which our extracted root differs from the root listed in the dictionary were noted on the wordlists.

2.2 Processing Wordlists

Here I describe the processes of IPA conversion, variant pruning, and tokenization, which serve as prerequisites for the next step of automatic cognate detection.

2.2.1 IPA Conversion

The first processing step was to convert the orthographic representation of roots in the wordlists into a uniform IPA representation. This was done by compiling the phonological inventories described in

Section 1.3 and applying the conversions in Table 12, using either a short Python script I wrote or the find-and-replace function in Google Sheets. Conversions involving longer spellings take precedent, which ensures that for example <ty> in Ashéninka becomes /tʃ/ rather than /tj/.

I do not use /N/ in my IPA representation because it would not be recognized as valid IPA by LingPy, which I used in the next step for automatic cognate detection. Instead, /N/ is always explicitly written with the place of articulation of the following consonant. In practice, this entailed doing a second stage of conversion in which /nk/ was changed to /ŋk/ and /ntʃ/ was changed to /ɲtʃ/. Arbitrarily, /nt^{j}/ was retained, even though I treat /t^{j}/ as alveopalatal. In any case, it is clear where /N/ needs to be reconstructed in Proto-Kampa, so its representation is of little importance.

In this step, I also removed tone marking from my Nomatsigenga data, for the same reason that vowels with acute accents are not part of standard IPA, so their behavior in LingPy was unclear to me. Tone marking has been reintroduced in the data presented in Section 4.

One ambiguity in the IPA representation arises from <Cia> sequences in Nomatsigenga. Specifically, Lawrence analyzes several words as having /Cja/ when they are written with <Cia> in Shaver's dictionary (Lawrence 2013, p. 14). However, it is clear that not all <Cia> sequences in Shaver's dictionary can be /Cja/, as the /i/ can carry high tone, like in /itsonía/ (Shaver 1996, p. 140). Since it is unclear to me in which cases <Cia> represents /Cja/ and whether Lawrence's analysis is correct, I have uniformly represented it as /Cia/.

2.2.2 Variant Pruning

The next step of wordlist processing was variant pruning, which entails reducing the number of forms for a given concept in a variety. This is beneficial because it cleans up the data in general and because according to Michael's personal conversation with List, the algorithm for cognate set construction described in Section 2.3 works best when there is an average of 1.5 roots per concept or fewer.

There is no formula for determining whether to retain or eliminate a form, but I was guided by the following general principles:

1. Keep multiple forms if they are true synonyms or pseudo-synonyms, which means that they have different meanings that fall under the same English concept.

2. Remove a form if it has a very specific meaning or one that is tangentially related to the concept, if a better form exists.

3. Remove a form if it is also listed for another concept which more closely matches the meaning.

4. For (morpho)phonological alternants, keep the form that is most likely to be the underlying one.

5. For sociolinguistic or dialectal variants, keep the form that is most common if such information is indicated. Otherwise, keep the form that appears to be the most conservative.[1]

6. When an excessive number of pseudo-synonyms exists (e.g. different species in the same genus), remove forms that have no cognates in the other Kampa varieties.

[1] Dictionaries cannot possibly capture and document all sociolinguistic and dialectal variation among speakers of a variety. As a result, despite these pruning efforts, my data will necessary contain forms which do not actually co-occur in any idiolect. This is especially true for Ashéninka, whose dictionary combines data from five distinct dialectal regions (Payne 1980, p. 4). Sociolinguistic and dialectal variation is likely the source of some apparent irregularities in the lexical data presented in Section 4.

grapheme	Nom	Ashé	Paj	Ashá	Kak	Mat
a	a	a	a	a	a	a
b	b					
c	k	k		k		
ch	tʃ	tʃʰ	tʃ	tʃ		tʃ
e	e	e	e	e	e	e
ë	ɨ					
g	g	ɣ	ɣ		ɣ	g
gu	g					
h					h	
i	i	i	i	i	i	i
j	h	h	h	h		h
k			k		k	k
ky			kʲ			kʲ
m	m	m	m	m	m	m
my		mʲ	mʲ			
n	n	n	n	n	n	n
ng	ŋ					
ngu	ŋ					
ñ		ɲ	ɲ			ɲ
ɲ					ɲ	
ŋ					ŋ	
o	o	o	o	o	o	o
p	p	p	p	p	p	p
py		pʲ	pʲ			
qu	k	k		k		
qy		kʲ				
r	ɾ	ɾ	ɾ	ɾ		ɾ
ry		ɾʲ	ɾʲ			
ɾ					ɾ	
s	s	s		s	s	s
sh	ʃ	ʃ	ʃ	ʃ		ʃ
ʃ					ʃ	
t	t	t	t	t	t	t
th		tʰ	tʰ			
ts	ts	tsʰ	tsʰ	ts	ts	ts
ty		tʃ	tʲ		tʲ	tʲ
tz		ts	ts			
tʃ					tʃ	
u						ɯi
v		β		β	β	b
β					β	
w			w			
wy			wʲ			
y	j	j	j	j	j	j

Table 12: Grapheme-IPA correspondences

2.2.3 Tokenization

The final step of wordlist processing was tokenization, which involves explicitly segmenting the IPA representation into words and phonemes. For example, /tsoɾoɾi haneki/ becomes /ts o ɾ o ɾ i _ h a n e k i/, where tokens are separated by spaces and the underscore indicates a word boundary. This was accomplished with a simple Python function which calls LingPy's ipa2tokens method (List, Greenhill, and Forkel 2017).

One caveat is that although I follow previous sources in analyzing orthographic double vowels as a single long vowel phoneme rather than a sequence of two identical vowels, I chose to tokenize them as two tokens. This is because in many cases, this results in a better alignment (alignments are discussed in more detail in the Section 2.3). For example, in Figure 4, Ashéninka and Pajonal /a-a/ align nicely with Kakinte and Matsigenka /aɾo/, where the dash represents a gap, but there is no appealing alignment of /aː/ with /aɾo/.

2.3 Cognate Set Construction and Alignment

After the previous processing steps, I combined the wordlists for the six varieties. This combined wordlist has 793 concepts, with 953 roots from Nomatsigenga, 750 from Asheninka, 1037 from Pajonal, 1051 from Ashaninka, 726 from Kakinte, and 923 from Matsigenka.

2.3.1 Automatic Cognate Detection and Alignment with LingPy

At this stage, I used LingPy, a Python library for historical linguistics developed by Johann-Mattis List, to automatically create cognate sets and alignments (List, Greenhill, and Forkel 2017). I relied on List's LingPy tutorial to guide me in writing the code (List 2017b). See this tutorial for a brief exposition about the algorithms described here, and see List's book on sequence comparison for an in-depth theoretical discussion (List 2014).

Computational tools for cognate detection and correspondence set construction were used solely to increase efficiency. All substantive judgments that were made algorithmically were checked and revised manually by my advisor and I, so ideally the use of computational tools had no effect on the reconstruction. In practice, it likely had a slight effect, as we probably tended to retain the algorithmic judgment if the correct judgment was unclear to us.

LingPy only finds cognates that share the same concept. Thus, cognate detection boils down to clustering roots for the same concept in various languages into cognate sets. LingPy provides several cognate detection algorithms, but I chose lexstat as it is the most reliable one when there are over 100 concepts in the data (List 2017b, p. 9). Briefly, the algorithm performs the following steps:

1. Convert the tokenized IPA representations into sound classes. I used the SCA system, which groups all IPA symbols into 25 sound classes. For example, /t/ and /tʃ/ are both converted into T. The purpose of this step is to "[reduce] variation which is unnecessary for historical comparison while at the same time maintaining a sufficient degree of distinctivity" (List 2017b, p. 3).

2. Generate a scoring function for each pair of languages which quantifies the similarity between potential cognates in those languages. This scoring function is "based on permutation statistics which give hints regarding the most likely candidates for regular sound correspondences" (List 2017b, p. 8).

3. Use the scoring functions to cluster words into cognate sets. I used the network-based infomap clustering algorithm, which is supposed to yield the best results (List 2017b, p. 11).

The next step in the LingPy pipeline is automatic alignment of cognates. An alignment is an explicit representation of corresponding sounds among cognates. A sample alignment is given below, using the root for "AFRAID" in five Kampa varieties. A dash indicates the absence of a sound.

Nom	ts	o	ɾ	o	g
Ashé	t^h	a	-	a	β
Paj	t^h	a	-	a	w
Kak	ts	a	ɾ	o	ɣ
Mat	ts	a	ɾ	o	g

Figure 4: Alignment example

List points out that although linguists have always implicitly aligned words, they have rarely used explicit alignments (List 2014, pp. 134–135). However, explicit alignments are necessary in LingPy because the columns of these alignments form the sites which are clustered into correspondence sets in the next step.

LingPy uses the pairwise scoring function generated in the cognate detection step as the basis for evaluating possible alignments (List 2017b, p. 11). Finding the optimal alignment of more than two sequences is computationally expensive, so LingPy provides progressive and library-based approximation algorithms for automatic alignment. I used the library-based algorithm, which works by first creating a library of optimal pairwise alignments, and then finding the overall alignment that is maximally consistent with this library (List 2014, pp. 108–109).

2.3.2 Editing Cognate Sets and Alignments with Edictor

I used Edictor, a browser-based application created by List, to edit the automatically generated cognate sets and alignments (List 2018a). See List's tutorial for an overview of Edictor's features (List 2017a). The changes made in this step fall into three categories:

1. I split polymorphemic forms into their constituent morphemes and created new cognate sets and alignments for each morpheme. List uses the term "partial cognates" to refer to these cognate morphemes (List 2017a, p. 29).

2. I combined cognate sets spanning multiple concepts into a single set. These cognate sets include morphemes found in different words as well as words that have undergone semantic shift.

3. I fixed errors made by LingPy, i.e. clustering non-cognates, failing to cluster cognates, and misaligning cognates. These judgments were made by my advisor and I drawing on our knowledge of the Kampa family, sound changes within the family, and general linguistic principles. I continually revised cognate sets and alignments in later steps as I figured out more about the relevant sound changes.

2.3.3 Dealing with Allomorphy

Ideally, in the context of the LingPy workflow, allomorphs should be placed in different cognate sets. This is because in the following step of correspondence set construction, a site (described in Section 2.4) is defined as having exactly one symbol for each variety.

If an instance of allomorphy is reconstructable to Proto-Kampa, then it is simple to separate the allomorphs in the daughter varieties. If allomorphy exists in only some varieties and is the predictable result of a sound change or phonological rule, then cognate sets should be created in accordance with the conditioning environment. For example, consider the morphemes /t^hapi/ in Ashéninka from /t^hapi-ki/, /t^hap^j/ in Ashéninka from /o-t^hap^j-a/, and /tsapi/ in Kakinte from /tsapi-ha/. In this case, /t^hap^j/ and /tsapi/ should be in the same cognate set because the correspondence of Ashéninka /p^j/ and Kakinte /p/ is regular in the PK environment **piha*, and /t^hapi/ should be in a different cognate set. Instances in which the cause of allomorphy is unclear should be handled on a case-by-case basis.

However, due to time limitations, I was unable to separate all instances of allomorphy into different cognate sets. In such cases, when a variety has more than one sound in a given site, LingRex

chooses one sound arbitrarily. This results in some complications in the later step of correspondence set naming, but as these cases are quite uncommon, for clarity I will not discuss this issue further.

2.4 Correspondence Set Construction

For automatic correspondence set construction, I used LingRex, a Python package written by Johann-Mattis List (List 2018b). Here I will briefly describe how LingRex's algorithm works in the context of my data. See List's paper for a more detailed discussion (List 2019).

The result of the previous stage is a collection of aligned cognates. Let a site denote a column of an alignment. For example, the alignment shown in Figure 4 consists of five sites. Each site comprises one symbol from each variety: a sound or gap if the variety is represented in the cognate set, and a null symbol otherwise.

Two sites are deemed compatible if the following two conditions hold: (a) for all varieties in which neither site has a null symbol, both sites have the same symbol, and (b) at least one of these shared symbols is not a gap. As an illustration, consider the three hypothetical sites shown below. Site 1 (the first row) and site 2 are compatible, as they agree for Nom, Paj, and Mat. Sites 2 and 3 are incompatible because they disagree for Ashá. Sites 1 and 3 are incompatible because although they agree for Ashé and Paj, both of these shared symbols are gaps. I hope this example shows why the definition of compatibility given above is sensible.

Nom	Ashé	Paj	Ashá	Kak	Mat
ɾ	-	-	∅	ɾ	ɾ
ɾ	∅	-	ɾ	∅	ɾ
∅	-	-	h	∅	∅

Figure 5: Demonstration of compatibility among sites

Given a collection of sites that are pairwise compatible, the non-null symbols (i.e. sounds and gaps) can be aggregated to form a correspondence set. For example, sites 1 and 2 above would induce the correspondence set /ɾ - - ɾ ɾ ɾ/. There are many ways to cluster all sites into correspondence sets, but it stands to reason that minimizing the total number of correspondence sets is generally preferable. Lingrex accomplishes this by reducing the problem to the minimum clique cover problem from graph theory.

LingRex outputs a three-part name for each site which identifies which correspondence set it was placed in. For example, the first site in the alignment shown in Figure 4 may be placed in the correspondence set 17-50/ts. The first number indicates the relative frequency of the correspondence set (with 1 being the most common), the second number indicates the number of sites in the set, and the final symbol is a sound or gap represented in the correspondence set (often, but not always, the most common symbol). A cognate set with only one form provides no information about correspondences, so sites for such cognate sets are assigned to the null correspondence set named 0/n.

2.5 Reconstruction

Finally, I used the correspondence sets generated by LingRex to reconstruct the consonant phonemes of Proto-Kampa and the sound changes that created the reflexes in the daughter varieties. During this step, I made extensive use of the correspondences panel in Edictor, which allowed me to easily view all instances of a correspondence set (List 2017a, pp. 37–38). I used the sound changes described by Michael as a starting point, but I did not assume that they were accurate. I followed standard procedures in the comparative method, using correspondence sets in complementary distribution as evidence of a common proto-phoneme, and using my knowledge of cross-linguistic sound changes to reconstruct the values of proto-phonemes.

Conveniently, I was able to keep track of the reconstructed proto-phonemes and sound changes by modifying the three-part names that LingRex generated for the correspondence sets. I used the first number as the unique ID of the correspondence set, creating a new ID whenever I created a new set. I used the final symbol to indicate my reconstructed Proto-Kampa phoneme. Finally, I replaced the second number with the environment which triggered sound changes in the daughter varieties, and also indicated unexpected reflexes.

The figure below shows the cognate set from Figure 4 along with the name that I gave to each site's correspondence set. The first set name, 17-_a,o/ts, indicates that I reconstruct **ts* in a *_a* or *_o* environment, from which the reflex of /t^{h}/ in Ashéninka and Pajonal is expected. The second set name indicates that I reconstruct **a* and that the reflex in Nomatsigenga is /o/, but I have not documented the environment as I am not concerned with vowels. The third set name indicates that I reconstruct **ɾ* in a *a_o* environment. However, a gap is not the expected reflex in Ashéninka and Pajonal, so I additionally write "LossAshePaj". The fourth name is unchanged, again because I did not examine vowels. The last set name indicates that I reconstruct **g* in a *a_o* or *o_a* environment, from which the reflexes in the daughter varieties are predictable.

Nom	ts	o	ɾ	o	g
Ashé	t^{h}	a	-	a	β
Paj	t^{h}	a	-	a	w
Kak	ts	a	ɾ	o	ɣ
Mat	ts	a	ɾ	o	g
correspondence sets	17-_a,o/ts	16-oNom/a	1076-a_o|LossAshePaj/ɾ	111-10/o	61-a_o,o_a/g

Figure 6: Example of renamed correspondence sets

In many cases, I found it necessary to revise LingRex's assignment of sites to correspondence sets. The case of /k/ in the four cognate sets shown below illustrates this well. The first row demonstrates the regular correspondence /k k k k k k/, which LingRex names 4-347/k and which I rename to 4-G/k (G signifies the most general environment). The second row demonstrates the correspondence /- k k k k k/ (the dash signifies the absence of a phoneme), which LingRex names 13-48/k and which I rename to 13-N_/k because **k* > ∅ / *N_* is a regular sound change in Nomatsigenga. The third cognate set is missing data from Nomatsigenga, so the site containing /k/ is compatible with both 4-347/k and 13-48/k. LingRex places it in 4-347/k, but the correct placement is with 13-N_/k, because we would expect a hypothetical Nomatsigenga cognate to have lost **k*. The last cognate set shows another instance of the correspondence /k k k k k k/, so naturally LingRex places the site containing /k/ in the correspondence set 4-347/k. However, the presence of /k/ in Nomatsigenga is actually unexpected in this environment. In order to capture this fact, I placed this site in a newly created correspondence set, which I named 1021-N_|RetainedNom/k.

	Nom	Ashé	Paj	Ashá	Kak	Mat	original set	revised set
PULL-2	nósik	noʃik	noʃik	noʃik	noʃik	noʃik	4-347/k	4-G/k
BONE	toŋi	toŋki	toŋki	toŋki	toŋki	toŋki	13-48/k	13-N_/k
SCRATCH (V.)				kitʃaŋk		kitʃaŋk	4-347/k	13-N_/k
Saccharum sp.	soŋko	saŋko	haŋko	saŋko	saŋko	ʃaŋko	4-347/k	1021-N_|RetainedNom/k

Table 13: Example of revising correpsondence sets

Like the discussion about allomorphy, the previous paragraph describes what would happen in the ideal case. In reality, time constraints prevented me from completely sorting all sites into their proper correspondence sets.

3 Proto-Kampa Consonant Inventory

In this section, I will lay out my reconstructed Proto-Kampa consonant inventory and use evidence from correspondence sets to argue for its correctness. In many cases, it is clear which Proto-Kampa phoneme should be reconstructed for a given correspondence set, but details about the sound changes that created the reflexes in the daughter varieties are unclear. These details have been suppressed in this section and are discussed in Section 4.

In the correspondence sets presented in this section, an empty cell signifies the absence of data and ∅ signifies the absence of a phoneme, i.e. the sound has been lost in that variety.

The following tables show my reconstruction of the Proto-Kampa consonant phonemes and my working model of the Proto-Kampa vowel phonemes. Before I begin my discussion of consonants, I will briefly elaborate on the vowel inventory. Short and long **i*,**e*, **a*, and **o* are tentatively reconstructed based on their presumed existence in all daughter varieties (Mihas 2017, p. 786). **ɯi* is reconstructed from the correspondence of Matsigenka /ɯi/, Nomatsigenga /ɨ/, and /i/ in the other varieties in my sample (see Table 72 for some examples). This PK phoneme is likely monomoraic as in the daughter varieties, but its precise phonetic realization is unclear. I follow Michael et al.'s reconstruction in calling it **ɯi* and classifying it as high and central (Michael et al. 2010, p. 8).

	labial	alveolar	(alveo)palatal	velar	glottal	unspecified
nasal	m	n				N
stop	p b	t		k g		
affricate		ts	tʃ			
fricative		s	ʃ		h	
liquid		ɾ				
glide			j			

Table 14: Reconstructed Proto-Kampa consonant inventory

	front	central	back
high	i iː	ɯi	
mid	e eː		o oː
low		a aː	

Table 15: Working model of Proto-Kampa vowel inventory

3.1 Reconstruction of **m*

I reconstruct PK **m* for the following correspondence sets. It is clear that *m must be reconstructed for the identity set, i.e. /m m m m m m/. The other set is only attested in one cognate set (see Section 4.1.1). In that instance, if the root can be reconstructed to Proto-Kampa at all, the PK root must contain the sequence **miha*, with the Pajonal reflex /mʲa/ developing from **miha* > *mia* > *mʲa*. The alternative reconstruction of **mʲ* cannot explain the reflex of /miha/ in Asháninka.

The phoneme /mʲ/ also exists in several other Pajonal words in my data which lack cognates. It is likely that these instances also developed from palatalization of a previous **m* at some stage in the history of the variety. It is far less likely that **mʲ* existed in Proto-Kampa and was lost in all varieties in my sample except Ashéninka and Pajonal.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
	mʲ	mʲ	m			*_iha*
m	m	m	m	m	m	elsewhere

Table 16: Correspondence sets for Proto-Kampa **m*

3.2 Reconstruction of **ɲ*

I reconstruct PK **ɲ* for the following correspondence sets. It is clear that **ɲ* must be reconstructed for the identity set. For the other set, I also reconstruct **ɲ* and posit the change **ɲ* > ɲ / *_ea, _ia, _io* in Ashéninka and Pajonal (see Section 4.1.2).

Alternatively, consider the reconstruction of **ɲ* for this set. First of all, this reconstruction cannot be correct for cases in which Kakinte has /niha/, because there is no reasonable pathway from **ɲa(a)* to /niha/ unless we invoke an unattested sound change of /h/-epenthesis. For the other cases, it is conceivable that PK **ɲ* was retained in Ashéninka and Pajonal and unpacked to /ne/ or /ni/ in the other varieties. Then, the development of /ne/ or /ni/ is unpredictable, unless one argues that only correspondences of /ɲ/ and /ni/ reflect PK **ɲ*. In the end, we are faced with the following facts: reconstructing only PK **ɲ* and positing palatalization to /ɲ/ in Ashéninka and Pajonal cleanly explains all of the data. Any reconstruction which includes **ɲ* must posit the cross-linguistically less common sound change of unpacking, and must also include the **ɲ* > ɲ change anyway. Thus, there is no reason to reconstruct **ɲ*.

This **ɲ* > ɲ change in Ashéninka and Pajonal is just one instance of a widespread palatalization sound change in these varieties (see Section 4.1). The argument above against reconstructing **ɲ* also applies to the other palatalized consonants in Ashéninka and Pajonal. Thus, I will refer to the above argument throughout this section without restating it in its entirety.

/ɲ/ is also a phoneme in Kakinte and Matsigenka (see Section 4.4.1). Again, it is conceivable that instances of /ɲ/ in Kakinte and Matsigenka reflect retentions of PK **ɲ*, but it is more likely that they result from a sporadic palatalization process like the regular process in Ashéninka and Pajonal. Consider the cognate set for "mosquito". If we reconstruct PK **maɲo*, then there is no good explanation for the vowel length discrepancy in Ashéninka /maɲoo/ and Pajonal /maɲo/. On the other hand, if we reconstruct **manio* or **maneo*, then we can explain the data by positing that regular palatalization with optional compensatory lengthening occurred in Ashéninka and Pajonal, while sporadic palatalization occurred in Matsigenka.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
n	ɲ	ɲ	n	n	n	*_ea, _eo, _i(h)a, _io*
n	n	n	n	n	n	elsewhere

Table 17: Correspondence sets for Proto-Kampa **n*

3.3 Reconstruction of **p*

I reconstruct PK **p* for the following correspondence sets. It is clear that **p* must be reconstructed for the identity set. I posit that the set /∅ p p p p p/ arose from the sound change **p* > ∅ / *N_* in Nomatsigenga (see Section 4.5.1). Alternatively, as pointed out by Michael, reconstructing **∅* for this set and positing post-nasal epenthesis in the non-Nomatsigenga varieties is flawed because there is no environment that predicts epenthesis (Michael et al. 2010, p. 3).

I also reconstruct **p* for the first correspondence set and posit the sound change **p* > *pʲ / _ea, _ia* in Ashéninka and Pajonal (see Section 4.1.3). The alternative analysis in which **pʲ* was unpacked in the other varieties is less likely (see the argument in Section 3.2 against the reconstruction of PK **ɲ*).

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
p	p^j	p^j	p	p	p	*_ea, _i(h)a*
∅	p	p	p	p	p	*N_*
p	p	p	p	p	p	elsewhere

Table 18: Correspondence sets for Proto-Kampa **p*

3.4 Reconstruction of **b*

I reconstruct **b* for the /b β w β β b/ correspondence set because lenition of **b* to /β/ is more common than fortition of **β* to /b/. Another (weaker) argument for **b* rather than **β* is that this phoneme serves as the voiced counterpart to **p* in the phonological rule of possessive voicing which is reconstructible to Proto-Kampa (see Section 2.1.1), so it makes more sense to reconstruct a stop.

I also reconstruct **b* for the first correspondence set and posit the sound change **w* > *w^j* / *_ea, _ia, _io* in Pajonal after **b* > *w* (see Section 4.1.4). See the argument against **n* in Section 3.2 for why I do not reconstruct a palatalized segment for this correspondence set. Note that in this case, reconstructing something like **b^j* is even less tenable because /$β^j$/ does not exist in Ashéninka.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
b	β, j	w^j	β	β	b	*_ea, _ia, _io*
b	β	w	β	β	b	elsewhere

Table 19: Correspondence sets for Proto-Kampa **b*

3.5 Reconstruction of **t*

I reconstruct PK **t* for the following correspondence sets. It is clear that **t* must be reconstructed for the identity set. For the set /t ts ts t t t/, I reconstruct **t* and posit the sound change **t* > *ts* / *_i* in Ashéninka and Pajonal (see Section 4.2.3). The alternative reconstruction of **ts* is unlikely because it would require the much rarer sound change of **ts* > *t* / *_i* in the other varieties.

I analyze Ashéninka /tʃ/ and Pajonal /t^j/ in the first two sets as coming from the palatalization of **t* before **ea*, **ia*, and **io* (see Section 4.1.5). In addition, I posit the sound change **t* > *t^j* / *_iV* in Matsigenka (see Section 4.7.1). This is complicated by the fact that there is no cognate set demonstrating the development of PK **tia* in Matsigenka, and the development of **tia* and **tio* in Kakinte is similarly unclear due to lack of cognates. In any case, **t* should be reconstructed in all of these instances for the same reasons outlined in Section 3.2, the most important of which being the improbability of **t^j* > *ti* unpacking.

The last two correspondence sets are less straightforward to reconstruct, as they are not in complementary distribution with the top four sets, suggesting the possibility of a different proto-phoneme. Nevertheless, I reconstruct **t* because there is simply no other reasonable option. For the /t ts^h ts t t t/ set, **t* must be reconstructed because something like **ts* or **ts^h* cannot explain the reflexes of /t/. I posit that /ts^h/ in Ashéninka arose from **t* unpredictably undergoing both stages of the **t* > *ts* > *ts^h* / *_i* chain shift, possibly because there was a period during which both **t* > *ts* and **ts* > *ts^h* were operative (see Section 4.2.3). For the /t t t t t^j t/ set, a reconstruction of **t^j* would require the unlikely merger of **t^j* and **t* elsewhere in the Kampa family. Instead, I suggest that **t* > *t^j* is sporadic in Kakinte, and is possibly a result of esoterogeny (see Section 4.7.2).

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
t	tʃ	tʲ	t	(tʲ)	tʲ	*_ia, _io*
t	tʃ	tʲ	t	t	t	*_e(h)a, _iha*
t	ts	ts	t	t	t	*_i*
t	t	t	t	t	t	elsewhere
t	tsʰ	ts	t	t	t	*_i*
t	t	t	t	tʲ	t	unclear

Table 20: Correspondence sets for Proto-Kampa **t*

3.6 Reconstruction of **k*

I reconstruct PK **k* for the following correspondence sets. It is clear that **k* must be reconstructed for the identity set. I posit that the set /∅ k k k k k/ arose from the sound change **k* > ∅ / *N_* in Nomatsigenga (see Section 4.8.1). Alternatively, as pointed out by Michael, reconstructing **∅* for this set and positing post-nasal epenthesis in the non-Nomatsigenga varieties is flawed because there is no environment that predicts epenthesis (Michael et al. 2010, p. 3).

I also reconstruct **k* for the first correspondence set and posit the sound change **k* > *kʲ / _ia* in Ashéninka and Pajonal (see Section 4.1.6). The alternative analysis in which **kʲ* was unpacked in the other varieties is less likely (see the argument in Section 3.2 against the reconstruction of PK **ɲ*).

/kʲ/ is also a phoneme in Matsigenka, but it is very rare and does not appear in any cognate sets in my data. Clearly, this does not provide a compelling reason to reconstruct **kʲ* for Proto-Kampa.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
k	kʲ	kʲ	k	k	k	*_i(h)a*
∅	k	k	k	k	k	*N_*
k	k	k	k	k	k	elsewhere

Table 21: Correspondence sets for Proto-Kampa **k*

3.7 Reconstruction of **g*

The following correspondence sets are in complementary distribution, with the caveat that I am combining two Kakinte reflexes into one correspondence set. This strongly suggests that all of these sets come from the same Proto-Kampa phoneme. I reconstruct **g* for all of these sets, and posit regular processes of lenition/loss in Ashéninka, Pajonal, and Asháninka, as well as irregular loss of **g* in Kakinte (see Section 4.9). Indeed, although details about the sound changes may be unclear, it is clear that **g* must be reconstructed in order to explain the reflexes of /g/ in Nomatsigenga and Matsigenka.

The only other plausible alternative is to uniformly reconstruct **ɣ* instead of **g*, but this is less likely as it necessitates the cross-linguistically rarer change of fortition in Nomatsigenga and Matsigenka. Another argument for PK **g* instead of **ɣ* is that this phoneme serves as the voiced counterpart to **k* in the morphophonological rule of possessive voicing, which can be reconstructed to Proto-Kampa (see Section 2.1.1).

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
g	j	j	j	ɣ, ∅	g	_*e*
g	β	w	∅	ɣ, ∅	g	*a*_*o*, *o*_*a*
g	ɣ	ɣ	∅	ɣ, ∅	g	*a*_#
g	j	j	∅	ɣ, ∅	g	*e*_*a*, *e*_*i*, *i*_a, *i*_*o*
g	∅	∅	∅	ɣ, ∅	g	elsewhere

Table 22: Correspondence sets for Proto-Kampa **g*

3.8 Reconstruction of **ts*

The following correspondence sets are in complementary distribution and include all possible **tsV* sequences except **tsui*, which does not appear in any good cognate sets in my data. This strongly suggests that all of these sets come from the same Proto-Kampa phoneme. I reconstruct **ts* for all of these sets and posit a three-way split of **ts* into /tʃʰ/ or /tʃ/, /tsʰ/, and /tʰ/ in Ashéninka and Pajonal (see Sections 4.2.1 and 4.1.7).

Now I will consider possible alternative reconstructions. For the first set, it would be conceivable to reconstruct **tʃ* or **tʃʰ*. However, this would require an alveopalatal consonant to become an alveolar or dental /ts/ before the front vowels /e/ and /i/, which is the opposite direction from the expected change.

For the second set, reconstructing **tsʰ* would mean that Proto-Kampa had **tsʰ* only before *i* and did not have **ts* before *i*. Then, there would be no basis to treat **ts* and **tsʰ* as separate phonemes in Proto-Kampa, and this analysis would amount to saying that this allophonic contrast became phonemic in Ashéninka and Pajonal with the introduction of /tsi/ but was lost in other Kampa varieties. This sequence of events is unnecessarily complicated. Similarly, reconstructing **tʰ* for the third set would essentially make **ts* and **tʰ* allophones in Proto-Kampa, and there is no good reason to do so.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
ts	tʃʰ	tʃ	ts	ts	ts	_*e*, _*ia*
ts	tsʰ	tsʰ	ts	ts	ts	_*i*
ts	tʰ	tʰ	ts	ts	ts	_*a*, _*o*

Table 23: Correspondence sets for Proto-Kampa **ts*

3.9 Reconstruction of **tʃ*

I trivially reconstruct **tʃ* for the following correspondence set. The alternative reconstruction of **tʃʰ* would make it the only aspirated consonant in Proto-Kampa and it would lack an unaspirated counterpart, so this is unlikely.

Heitzman does not include **tʃ* in her reconstruction, and instead argues that /tʃa/ and /tʃo/ sequences in all daughter varieties come from Proto-Kampa **tʰia* and **tʰio* (Heitzman 1973, p. 13). First of all, I do not reconstruct PK **tʰ* at all, but I will entertain the alternative hypothesis that /tʃa/ and /tʃo/ come from **tsia* and **tsio*. This is problematic because it does not explain the origin of /tsia/ sequences (e.g. in /katsiaɾi/ meaning "SMOKE (n.)") in a variety like Nomatsigenga, which in my analysis simply come from **tsia*.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
tʃ	tʃʰ	tʃ	tʃ	tʃ	tʃ	_*a*, _*o*

Table 24: Correspondence set for Proto-Kampa **tʃ*

3.10 Reconstruction of **s*

I reconstruct **s* for the following correspondence sets. Clearly, **s* must be reconstructed for the /s s h s s s/ set, because the alternative reconstruction of **h* would require an unlikely **s* > *h* sound change. I propose that the first two correspondence sets result from **s* > ʃ / _*i* in every variety except Nomatsigenga, and that this change was fed by **ui* > *i* in Ashéninka, Pajonal, Asháninka, and Kakinte (see Section 4.12.1). Also, I posit **s* > ʃ / _*e* in Ashéninka, Pajonal, and Asháninka, as well as occasionally in Kakinte (see Section 4.12.2). The alternative reconstruction of **ʃ* for the first or third set would necessitate an unlikely change of **ʃ* to */s/* before a front vowel. Reconstructing **ʃ* for the second set would require **ʃ* > *s* in Nomatsigenga and Matsigenka before *ui* but not before *a* or *o*, and there is no phonetic motivation for this change.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
s	ʃ	ʃ	ʃ	ʃ	ʃ	_*i*
s	ʃ	ʃ	ʃ	ʃ	s	_*ui*
s	ʃ	ʃ	ʃ	s, ʃ	s	_*e*
s	s	h	s	s	s	_*a*, _*o*

Table 25: Correspondence sets for Proto-Kampa **s*

3.11 Reconstruction of **ʃ*

I trivially reconstruct **ʃ* for the following correspondence set. Heitzman does not include **ʃ* in her reconstruction, and instead argues that /ʃa/ and /ʃo/ sequences in all daughter varieties come from Proto-Kampa **sia* and **sio* (Heitzman 1973, p. 13). This analysis is problematic because it does not explain the origin of /sia/ sequences in a variety like Nomatsigenga (e.g. /siá/ meaning "MYRMECOPHAGA TRIDACTYLA"), which in my analysis simply come from **sia*.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
ʃ	ʃ	ʃ	ʃ	ʃ	ʃ	_*a*, _*o*

Table 26: Correspondence set for Proto-Kampa **ʃ*

3.12 Reconstruction of **h*

There are numerous correspondence sets made of /h/ and ∅, several of which are shown below. In general, these sets are not in complementary distribution. However, in all of these cases it is clear that we must reconstruct PK **h* and admit that the conditions for **h*-loss are not well understood (see Section 4.14 for a discussion of some discernible patterns). The only other reasonable alternative is to reconstruct **∅ in some cases and posit epenthesis of /h/, but this is a rarer sound change than **h*-loss and it does not seem to occur in any regular environment.

Nom	Ashé	Paj	Ashá	Kak	Mat
h	h	∅	h		h
h	h	h	h	h	∅
h	∅	∅	h	h	∅
∅	∅	∅	h	h	∅
∅	∅	∅	∅	h	∅

Table 27: Correspondence sets for Proto-Kampa **h*

3.13 Reconstruction of **ɾ*

I reconstruct **ɾ* for the following correspondence sets. Clearly, **ɾ* must be reconstructed for the identity set. For the /ɾ ɾʲ ɾʲ ɾ ɾ ɾ/ set, I reconstruct **ɾ* and posit the sound change **ɾ > ɾʲ / _ea, _ia* in Ashéninka and Pajonal (see Section 4.1.9). See Section 3.2 for my argument against reconstructing **ɲ*, which also applies to **ɾʲ*. For the second and third cognate sets, it is clear that **ɾ* must be reconstructed and that **ɾ* was lost in Pajonal between identical non-front vowels and became /w/ in an *a_o* environment (see Section 4.15.1).

The last correspondence set exists in the same environment as the third set. The only reasonable reconstruction is **ɾ*, as the alternative would require **∅* or some unattested weaker form of **ɾ* merging with **ɾ* in Nomatsigenga, Asháninka, Kakinte, and Matsigenka. I suggest that this correspondence set is most likely explained by the sampling of different dialects in the Ashéninka dictionary (see Section 4.15.2).

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
ɾ	ɾʲ	ɾʲ	ɾ	ɾ	ɾ	*_ea, _i(h)a*
ɾ	ɾ	w, ∅	ɾ	ɾ	ɾ	*a_o*
ɾ	ɾ	∅	ɾ	ɾ	ɾ	*a_a, o_o*
ɾ	ɾ	ɾ	ɾ	ɾ	ɾ	elsewhere
ɾ	∅	∅	ˊɾ	ɾ	ɾ	*a_a, o_o*

Table 28: Correspondence sets for Proto-Kampa **ɾ*

3.14 Reconstruction of **j*

I trivially reconstruct **j* for the following correspondence set.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
j	j	j	j	j	j	*_a, _o*

Table 29: Correspondence set for Proto-Kampa **j*

3.15 Reconstruction of **N*

I reconstruct **N* for the following correspondence sets. With the exception of instances of **Np > m* (see Section 4.5.1) and **Nk > ŋ* (see Section 4.8.1) in Nomatsigenga, all of these nasals are synchronically analyzed as /N/ in the daughter varieties. However, they surface with the place of articulation of the following consonant, and I write them in their surface form.

Basically, we can reconstruct surface [mp], [nt], [nts], [ɲtʃ], and [ŋk] sequences to Proto-Kampa. The same arguments for analyzing the nasals in these sequences as /N/ in the daughter varieties (see Section 1.3.1) apply to proto-Kampa. Thus, I reconstruct **N* in Proto-Kampa.

Nom	Ashé	Paj	Ashá	Kak	Mat	PK environment
m	m	m	m	m	m	*_p*
n	n	n	n	n	n	*_t, _ts*
n		ɲ			n	*_tsia*
ɲ	ɲ	ɲ	ɲ	ɲ	ɲ	*_tʃ*
ŋ	ŋ	ŋ	ŋ	ŋ	ŋ	*_k*

Table 30: Correspondence sets for Proto-Kampa **N*

3.16 Examination of Reconstructed Inventory

My reconstructed Proto-Kampa consonant inventory is quite typical from a cross-linguistic perspective. With the exception of the placeless nasal, which has already been discussed extensively, the only irregularity is that **d* is missing from the collection of stops. One might be inclined to analyze **ɾ* as underlying **d* to fill this gap, but Lawrence notes that in Nomatsigenga /t/-initial nominal roots do not undergo the possessive voicing rule described in Section 2.1.1, so /ɾ/ does not serve as the voiced counterpart of /t/ (Lawrence 2013, p. 9). The same argument applies to Proto-Kampa.

Figure 7 shows the frequency of consonant phonemes among 44 Arawak varieties, nine of which are Kampa. The boxed phonemes are the ones that I reconstruct for Proto-Kampa. We see that all of my reconstructed phonemes are quite common in the Arawak family except for /N/ and /g/. All common consonant phonemes are in my reconstruction except for /w/. Overall, my reconstructed inventory matches the general Arawak profile quite well. This suggests that it is reasonable but obviously does not serve as positive evidence for its accuracy.

4 Sound Changes

In this section, I lay out the sound changes affecting Proto-Kampa consonants in the six daughter varieties and show how they resulted in their diverse phonemic inventories. I first describe two processes in Ashéninka and Pajonal that affected multiple consonants: palatalization and the development of contrastive aspiration. I then go through each Proto-Kampa consonant and describe the other sound changes that affected it. Finally, I summarize the sound changes that occurred in the history of each variety.

In the cognate sets presented in this section, the given gloss reflects the most common meaning among the represented varieties. It should not be interpreted as necessarily reflecting the meaning in all Kampa varieties or the meaning of the Proto-Kampa form. In cases of polysemy (e.g. Kampa varieties often use the same words for "DAUGHTER-IN-LAW" and "NIECE"), one gloss is chosen arbitrarily.

The forms shown in the cognate sets are phonemic representations, with the exceptions that /N/ is written in its surface place-assimilated form and long vowels are written as double vowels. In general, the cognate sets show the full extracted form rather than just the cognate morpheme to more clearly show the complete phonetic environment of the relevant sound.

4.1 Palatalization in Ashéninka and Pajonal

In Ashéninka and Pajonal, consonants were generally palatalized before the sequences *ea*, *ia*, and *io*. That is, **Cea* and **Cia* became /C^{j}a/ (or /C^{j}aa/ when optional compensatory lengthening occurred), and **Cio* became /C^{j}o/ (or /C^{j}oo/ with compensatory lengthening). The conditions for compensatory lengthening are unclear to me, and as it is a vocalic phenomenon, I will not investigate it further. The situation before *ie* is murky, as it is a very rare sequence.

The palatalized reflexes of PK **ts*, **s*, and **∅* (see Section 4.1.10 for an explanation of **∅ > j*) merged with existing phonemes in Ashéninka and Pajonal, and the other palatalized consonants became new phonemes. This sound change was often fed by the loss of intervocalic **h*, e.g. **Ciha > Cia > C^{j}a*.

Table 31 summarizes the palatalized reflexes of each Proto-Kampa consonant in Ashéninka and Pajonal. **b* does not palatalize in Ashéninka; this is discussed in Section 4.1.4.

There are no examples of PK **tʃ*, **ʃ*, **h*, or **j* in an environment that would lead to palatalization in Ashéninka or Pajonal. Thus, these consonants are omitted from the table and the following discussion. I also omit **g* because it is unclear whether its lenition and loss, described in Section 4.9.2, occurred before or after this palatalization. See Section 4.1.10 for examples of interactions between palatalization and **g*-loss.

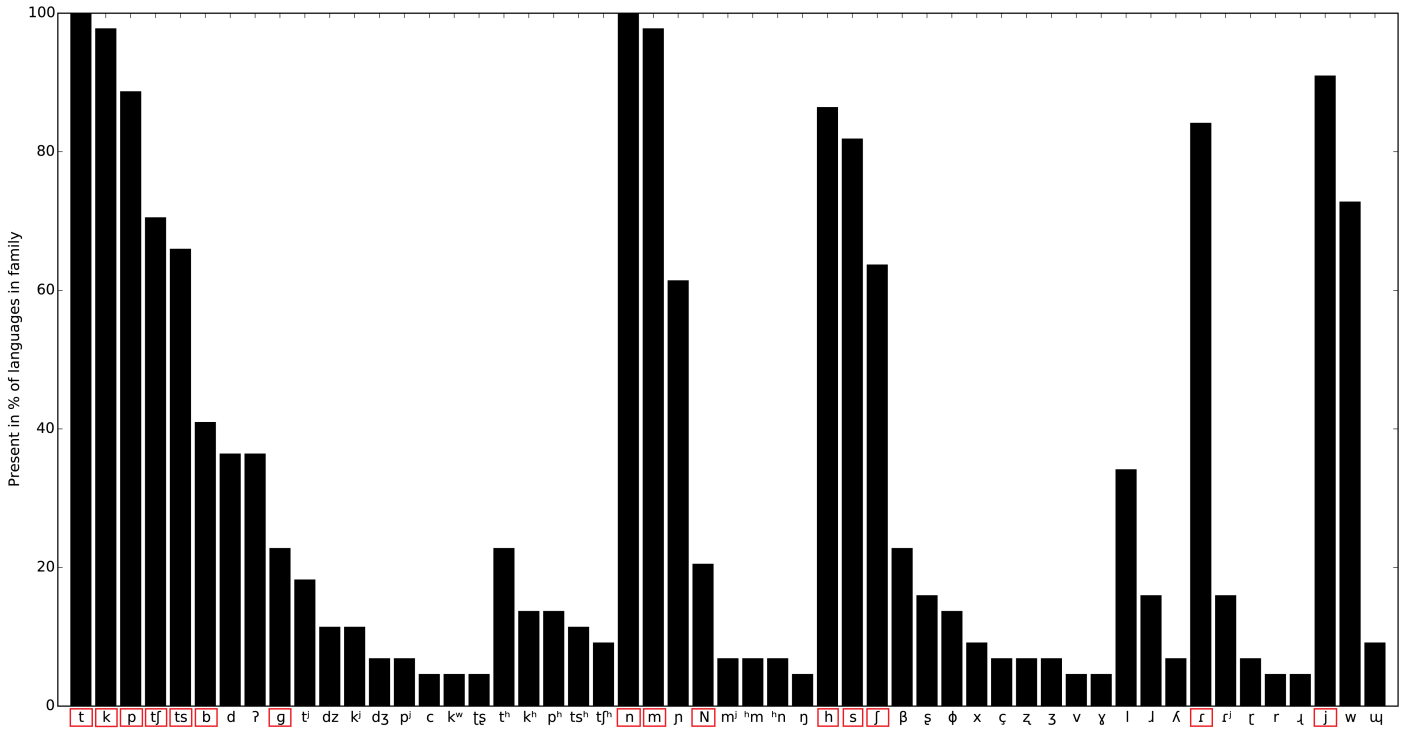

Figure 7: Consonant phoneme frequencies among Arawak varieties (Clem and Michael 2015)

PK	Ashé	Paj
**m*	mʲ	mʲ
**n*	ɲ	ɲ
**p*	pʲ	pʲ
**b*		wʲ
**t*	tʃ	tʲ
**k*	kʲ	kʲ
**ts*	tʃʰ	tʃ
**s*	ʃ	ʃ
**ɾ*	ɾʲ	ɾʲ
**∅*	j	j

Table 31: Palatalized reflexes of Proto-Kampa consonants in Ashéninka and Pajonal

Payne's book about the Apurucayali dialect of Ashéninka describes a similar synchronic phonological rule of palatalization, namely /Cia/ → [Cʲa] (Payne 1981, p. 128). If we accept his analysis, then several instances were I have posited the diachronic change **C* > *Cʲ* (e.g. **ʃiŋkia* > *ʃiŋkʲa* in Pajonal) are better analyzed as examples of this phonological rule (e.g. /ʃiŋki-a/ → [ʃiŋkʲa]). In any case, this does not take away from the fact that the diachronic sound change of palatalization must have also occurred. Indeed, if we posit that all instances of [Cʲ] are underlying /Ci/ sequences, then it is difficult to explain compensatory lengthening. For example, surface [Cʲa] and [Cʲaa] seem to occur in the same environments, so they should have different underlying forms. If [Cʲa] is /Cia/, then [Cʲaa] would have to be /Ciaa/, which would require an unmotivated change of **Cia* > *Ciaa*.

There are also several differences between Payne's description and my description that can be attributed to dialectal differences. In Apurucayali Ashéninka, only apical and velar consonants have palatalized counterparts, while labial consonants instead trigger epenthesis of [j] to break up the /ia/ sequence, i.e. /Cia/ → [Cija]. In the variety of Ashéninka in my data, only /b/ lacks a palatalized counterpart. Also, in Apurucayali Ashéninka the palatalized counterparts of both /ts/ and /k/ are /tʃ/, whereas in my data they are /tʃʰ/ and /kʲ/ respectively.

In several cases, **Ce* or **Ci* sequences at the end of verb roots undergo palatalization to /Cʲ/ despite the root not containing the full palatalizing environment. In all of these cases, if a Kakinte cognate exists, it ends with /Ceh/ or /Cih/. PK **h* is generally retained in Kakinte and lost in Ashéninka and Pajonal (see Section 4.14). Verb roots in Ashéninka and Pajonal are often followed by /a/-initial suffixes (see (Payne 1981, pp. 234–237) for verbal paradigms).

These facts lead me to the following hypothesis. In the history of Ashéninka and Pajonal, PK **Ceh-a* became *Ce-a* due to loss of **h* (ditto for **Cih-a*), which palatalized to /Cʲ-a/. Then, root-final /Cʲ/ was generalized to the rest of the verbal paradigm. PK roots ending in **Ce* or **Ci* were always followed by an epenthetic **t* in the parallel forms **Ce-t-a* and **Ci-t-a*, so they never had the opportunity to palatalize. Time constraints prevented me from further investigating this hypothesis and looking for counterexamples.

4.1.1 **m* > *mʲ*

Palatalization of **m* in Ashéninka and Pajonal created the phoneme /mʲ/. I found only one cognate set showing this change. The roots /mʲ/ and /omʲaɣ/ both exist in Pajonal; the historical and synchronic morphological relationship between them is unclear.

	Nom	Ashé	Paj	Ashá	Kak	Mat
JUMP-1		mʲ	mʲ, omʲaɣ	miha		

Table 32: **m* > *mʲ / _ia* in Ashéninka and Pajonal

4.1.2 **n* > *ɲ*

Palatalization of **n* in Ashéninka and Pajonal created the phoneme /ɲ/. In the first two cognate sets, the root does not contain the complete palatalizing environment, and the Kakinte cognate /neh/ is /h/-final in accordance with my hypothesis.

The **n* > *ɲ* change in Matsigenka /maɲo/ and Kakinte /pooɲa/ is discussed in Section 4.4.1.

	Nom	Ashé	Paj	Ashá	Kak	Mat
SEE	nía	ɲ	ɲ	ne	neh	ne
LIVE		aɲ	aɲ	ane		
SPEAK		ɲaa	ɲaa	nea		ni
CACICUS SP.	poonia	pooɲa	pooɲa		pooɲa	
APPEAR		koɲaa	koɲaa		koniha	konea
WATER	nihá	ɲaa	ɲaa	niha	niha	nia
SHOW (V.)	oniág	oɲaak	oɲaaɣ	oniha	onihaɣ	
BROTHER-IN-LAW			ɲani		aniani	
mosquito	monio	maɲoo	maɲo	maneo		maɲo
NEPHEW-1		ɲotsʰi	ɲotsʰi	nioti		

Table 33: **n* > *ɲ* / *_ea, _ia, _io* in Ashéninka and Pajonal

4.1.3 **p* > *pʲ*

Palatalization of **p* in Ashéninka and Pajonal created the phoneme /pʲ/. The Ashéninka word /pijaɾe/ shows insertion of an epenthetic /j/ rather than palatalization. This could be because the word was taken from a variety without /pʲ/ like Apurucayali, in which case we expect such an epenthetic /j/ to break up the /ia/ sequence (Payne 1981, p. 128).

	Nom	Ashé	Paj	Ashá	Kak	Mat
WIND		tampʲaa	tampʲaa	tampea		tampia
chica, manioc	piárí	pijaɾe	pʲaaɾe	peaɾe		piaɾi
BOW-1	pía	pʲamine	pʲampeni			pia
shore, bank		tʰapʲa	tʰapʲa		tsapiha	tsapia

Table 34: **p* > *pʲ* / *_ea, _ia* in Ashéninka and Pajonal

4.1.4 **b* > *wʲ* in Pajonal

In Pajonal, palatalization of **w* from PK **b* created the phoneme /wʲ/ (**b* > *w* is discussed in Section 4.6). On the other hand, Ashéninka does not have the phoneme /βʲ/ (/βʲ/ appears in Payne's dictionary only once, in the word /kitesaβʲaɾi/ in the PI-Nevati dialect (Payne 1980, p. 116)). /βijoŋkaɾe/ shows the expected outcome, in which an epenthetic /j/ was inserted to break up the /io/ sequence. However, in /jaane/, the reflex of PK **b* in a palatalizing environment is /j/.

	Nom	Ashé	Paj	Ashá	Kak	Mat
BOW-2	bíane	jaane	wʲani	βeane	βeane	
BRIDGE	pobí		pawʲa	paβiha		pabi
LUNGS	boŋaɾe	βijoŋkaɾe	wʲoŋkaɾe	βihoŋkaɾi		βɯioŋkaɾe

Table 35: **w* > *wʲ* / *_ea, _ia, _io* in Pajonal

4.1.5 **t* > **tʃ* in Ashéninka, **t* > **tʲ* in Pajonal

In Pajonal, palatalization of **t* created the phoneme /tʲ/. In Ashéninka, palatalization of **t* recreated the phoneme /tʃ/, as PK **tʃ* had uniformly turned into /tʃʰ/ (see Figure 8). The palatalization shown in the Matsigenka word /pantʲo/ is discussed in Section 4.7.1.

In the Pajonal roots /ontʲaɣ/ and /montʲaɣ/, the palatalizing environment seems to have been created by the unexplained insertion of /aɣ/. The palatalization of /moɲtʃ/ in Ashéninka is expected according to my hypothesis, as the Kakinte cognate is /h/-final.

	Nom	Ashé	Paj	Ashá	Kak	Mat
FLEA			mentʲaki	penteaki	mantaki	
cross river (v.)	montia	moɲtʃ	montʲaɣ	monte	monteh	monte
OLD	antiasípá	antʃaʃipa	antʲaʃipa	antiaʃipa		
WEAVE			ontʲaɣ	tih	ontih	
STOMACH			motʲaa	motiha	motiha	motia
BURY	tía		tʲaaki	tih	tihaβio	
DAUGHTER		ʃiɲtʃo	ʃintʲo	ʃintio		
NEPHEW-2		tʃomi		tiomi		
DUCK		paɲtʃo	pantʲo	pantio		pantʲo

Table 36: **t* > **tʃ / _ea, _ia, _io* in Ashéninka, **t* > **tʲ / _ea, _ia, _io* in Pajonal

4.1.6 **k* > **kʲ*

Palatalization of **k* in Ashéninka and Pajonal created the phoneme /kʲ/. In the first three cognate sets, the root does not contain the complete palatalizing environment, and Asháninka and Kakinte have /h/-final roots in accordance with my hypothesis (root-final **h* is actually generally lost in Asháninka; see Section 4.14.2).

	Nom	Ashé	Paj	Ashá	Kak	Mat
ENTER	k	kʲ	kʲ	kia	kih	ki
carry on back	kía	kʲ	kʲaatʰa	kih		ki
LEAVE			hokʲ	sokih		
CHICHA, CORN	síŋiá		ʃiŋkʲa			ʃiŋkiato

Table 37: **k* > **kʲ / _ia* in Ashéninka and Pajonal

4.1.7 **ts* > **tʃʰ* in Ashéninka, **ts* > **tʃ* in Pajonal

In Pajonal, **ts* palatalized into /tʃ/, merging with the existing phoneme /tʃ/ as well as the reflex of PK **ts* before *e* (see Figure 9). In Ashéninka, based on just one example, it appears that **ts* palatalized into /tʃʰ/, also merging with the reflex of PK **ts* before *e* and PK **tʃ* (see Figure 8). It is unclear if the palatalization evident in the Asháninka form /katʃaaɾi/ results from a similar sound change or borrowing.

	Nom	Ashé	Paj	Ashá	Kak	Mat
CURRENT	sintsiá		ʃiɲtʃaa			ʃintsia
SMOKE	kátsiáɾi	katʃaaɾi	katʃaaɾe	katʃaaɾi		

Table 38: **ts* > **tʃʰ / _ia* in Ashéninka, **ts* > **tʃ / _ia* in Pajonal

4.1.8 **s* > *ʃ*

In Ashéninka and Pajonal (as well as Asháninka, Kakinte, and Matsigenka), PK **s* went to /ʃ/ before *i* (see Section 4.12.1). There is no evidence to indicate whether this change happened before or after the general palatalization that is the subject of the current subsection. If the general palatalization happened first, then we simply have **sia* > *ʃaa* in one step. Otherwise, we have **sia* > *ʃia* > *ʃaa*, in which case the second palatalization causes compensatory lengthening but does not actually change the already palatal consonant /ʃ/.

	Nom	Ashé	Paj	Ashá	Kak	Mat
MYRMECOPHAGA TRIDACTYLA	siá		ʃaa	ʃaa		ʃia

Table 39: **s* > *ʃ* / *_ia* in Ashéninka and Pajonal

4.1.9 **ɾ* > *ɾʲ*

Palatalization of **ɾ* in Ashéninka and Pajonal created the phoneme /ɾʲ/. In the first three cognate sets, the root (or the reversative suffix) does not contain the complete palatalizing environment, and the Kakinte cognates are /h/-final in accordance with my hypothesis.

	Nom	Ashé	Paj	Ashá	Kak	Mat
lie (on surface)	naɾiá	noɾʲ		naɾe	noɾih	noɾi
reversative	ɾia	ɾʲ	ɾʲ	ɾe	ɾeh	ɾe
THINK	síɾe	kiŋkiʃiɾʲ	keŋkitʰaʃiɾʲ	keŋkeʃiɾe		sɯiɾe
UNRIPE		natsʰiɾʲaa	natsʰiɾʲaa	atsiɾea	etsiɾia	
REST (vi.)			omaɾʲaɣ	moɾea		
DRY (vi.)-1	piɾiá	piɾʲaa	piɾʲaa	piɾihatahi	piɾih	piɾia
SMALL-1		iɾʲani		oɾihani	iɾihani	
FISH (V.)	siɾiá		ʃiɾʲat			ʃiɾia

Table 40: **ɾ* > *ɾʲ* / *_ea, _ia* in Ashéninka and Pajonal

4.1.10 **∅* > *j*

In Ashéninka and Pajonal, the sequences **ea* and **ia* do not exist. We have seen that when preceded by a consonant, these sequences led to /Cʲa/ or /Cʲaa/. When they occurred at the beginning of a word or preceded by a vowel, they became /ja/ or /jaa/. I propose that this change can analyzed in the same way as the consonant palatalizations if we view it as **∅ia* > *∅ʲa*, which is realized as /ja/ (with possible compensatory lengthening). In the cognate sets for "WAIT" and "BOIL (v.)", the root does not contain the complete palatalizing environment, and Asháninka and Kakinte have /h/-final cognates in accordance with my hypothesis.

The cognate sets for "WAIT" and "FOLLOW" show evidence of palatalization, **h*-loss, and **g*-loss (see Section 4.9.2 for discussion of **g*-loss). The ordering of these changes is unclear and merits further study. One possible history for Ashéninka and Pajonal /ojaa/ is **ogiha* > *oiha* > *oia* > *ojaa*, with the last stage being **∅* > *j*. However, a sequence like **ogiha* > *ogia* > *ogʲaa* > *ojaa* is also plausible.

The change from **oia* to /oja/ in the Asháninka root for "WAIT" is a separate change, which occurred because /ViV/ sequences are disallowed. The sequence /ia/ does exist in Asháninka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
BEE	eíróto		jaaɾato		eaɾoto	
WAIT	ogia	oj	ojaɣ	oja	oɣih	ogi
follow	agiá	ojaa	ojaa		oɣiha	ogia
BOIL (v.)		moj	moj	moih	moiha	

Table 41: *∅ > *j* / _*ea*, _*ia* in Ashéninka and Pajonal

4.2 Contrastive Aspiration in Ashéninka and Pajonal

Figure 8 shows the development of PK **t*, **ts*, and **tʃ* in Ashéninka. We see that two stages of sound changes occurred. In the first stage, aspirated consonants were created by the three-way split of **ts* into /tʰ/, /tsʰ/, and /tʃʰ/ and the uniform change of **tʃ* into /tʃʰ/. In the second stage, unaspirated /ts/ was recreated by the palatalization of **t* before *i* and unaspirated /tʃ/ was recreated by the palatalization of **t* before *ea*, *ia*, and *io* as described in Section 4.1.5. The two stages must be ordered in this way because **ti* generally does not become /tsʰi/ (exceptions are given below) and **t* never becomes /tʃʰ/. The result of these sound changes is that Ashéninka now has contrastive aspiration between the phoneme pairs /t/-/tʰ/, /ts/-/tsʰ/, and /tʃ/-/tʃʰ/.

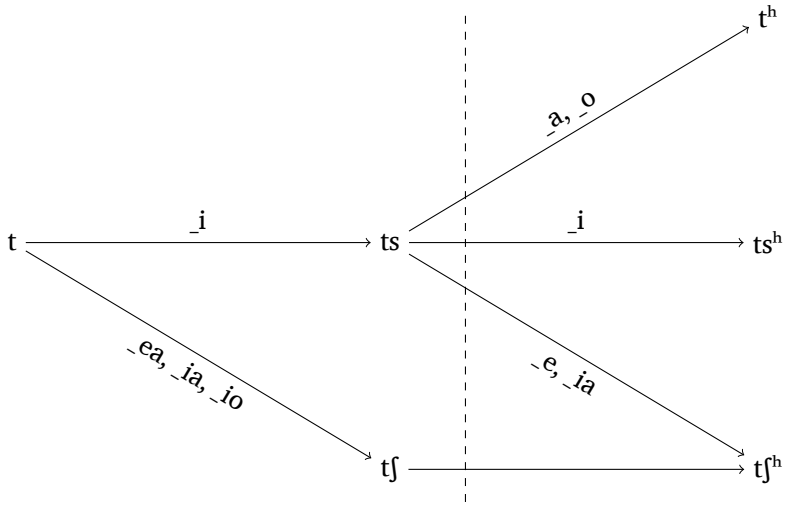

Figure 8: Development of **t*, **ts*, and **tʃ* in Ashéninka

One caveat is that it is unclear whether all six consonants should be analyzed as separate phonemes. [t] and [ts] are almost in complementary distribution, as in my data [ts] only occurs before /i/ and [t] does not occur before /i/ except in the word /iʃaaβiitiini/. Also, [tʰ] and [tsʰ] are in complementary distribution, as in my data [tʰ] only occurs before /a/ and /o/ and [tsʰ] only occurs before /i/. Thus, a conceivable analysis is that only /t/, /tʰ/, /tʃ/, and /tʃʰ/ are phonemes, and that [ts] is an allophone of /t/ and [tsʰ] is an allophone of /tʰ/. Under this analysis, Ashéninka would have two pairs of phonemes with contrastive aspiration rather than three. However, Payne's book about Apurucayali Ashéninka does list all six sounds as phonemes (Payne 1981, p. 59). I cannot find any argumentaion for this, but for simplicity I will assume that his analysis is correct for the remainder of the discussion.

Figure 9 shows the development of PK **t*, **ts*, and **tʃ* in Pajonal. Like in Ashéninka, the first stage of sound change saw a three-way split of **ts*, creating the aspirated phonemes /tʰ/ and /tsʰ/. However, **ts* before *e* and *ia* as well as **tʃ* developed into unaspirated /tʃ/. In the second stage, **t*

became /ts/ before *i* as in Ashéninka, but **t* before *ea*, *ia*, and *io* became /tʲ/. The result of these sounds changes is that Pajonal has two pairs of phonemes with contrastive aspiration: /t/-/tʰ/ and /ts/-/tsʰ/. Instead of the /tʃ/-/tʃʰ/ contrast in Ashéninka, it has a contrast between /tʲ/ and /tʃ/. The caveat about whether /ts/ and /tsʰ/ should actually be analyzed as phonemes also applies to Pajonal.

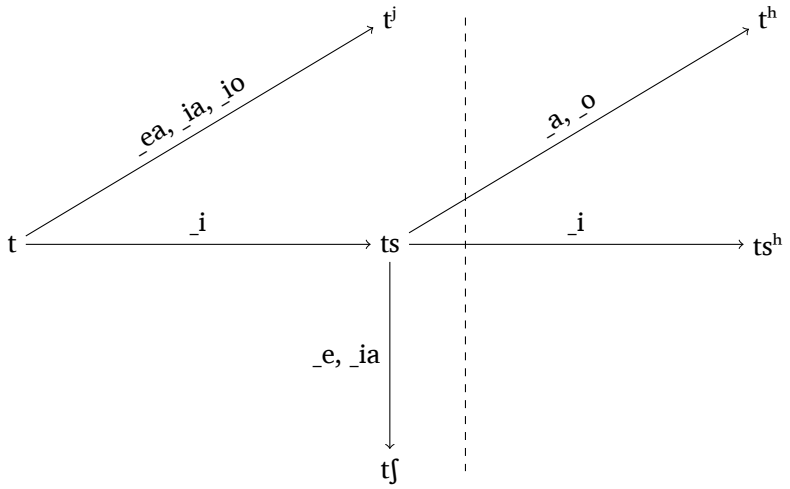

Figure 9: Development of **t*, **ts*, and **tʃ* in Pajonal

The development of contrastive aspiration in Ashéninka was previously noted by Michael et al. (Michael et al. 2010, p. 2). In the variety that they describe, **tʃ* did not become aspirated, so their description matches my description of Pajonal but not of Ashéninka.

According to Mihas, the development of contrastive aspiration in Ashéninka and Pajonal can possibly be attributed to contact with Quechua languages with contrastive aspiration (Mihas 2017, p. 786). I am not in a position to make any claim about the plausibility of this suggestion.

4.2.1 Three-way Split of **ts* in Ashéninka and Pajonal

The tables below show that in Ashéninka and Pajonal, PK **ts* became /tsʰ/ before *i* and /tʰ/ before *a* and *o*. Before *e*, PK **ts* became /tʃʰ/ in Ashéninka while it turned into /tʃ/ in Pajonal. I have no explanation for why the first **ts* in the cognate set for "HIGH, tall" became /s/ in Ashéninka and /h/ in Pajonal.

The Pajonal form /ʃiɲtʃeeŋka/, meaning "SHOUT, yell", provides evidence that **ts* > *tʃ* / _*e* occurred after **h*-loss in Pajonal. If we accept the etymology that it derives from PK **sintsi-heŋka* meaning "strong + air classifier", then the only reasonable sequence of sound changes at the morpheme boundary is **tsihe* > *tsee* > *tʃee*.

	Nom	Ashé	Paj	Ashá	Kak	Mat
COMPANION, friend	tsipa	tsʰipa	tsʰipataɾiɾi	tsipa	tsipaβihitaka	tsipa
INGA SP.	antsípá	intsʰipa	intsʰipa	intsipa	intsipa	intsipa
PUSH	otatsiŋ	tatsʰiŋk	otatsʰiŋk	tatsiŋk	tatsiŋk	tatsiŋk
Saimiri sp.	tsigiɾi	tsʰijeɾi	tsʰijeɾi	tsijeɾi	tsieɾi	tsigeɾi
earth, dirt, soil, ground	kibatsi	kipatsʰi	kipatsʰi	kipatsi	kepatsi	kipatsi
HURT, be painful (vi.)	katsi	katsʰi	katsʰi	katsi	katsi	katsi
BIRD (generic)	tsímiɾi	tsʰimeɾi	tsʰimeɾi	tsimeɾi	tsimeɾi	tsimeɾi
BITE (V.)	atsík	atsʰik	atsʰik	atsi	atsik	atsik
COLD (ADJ.)	kátsiŋá	katsʰiŋka	katsʰiŋka	katsiiŋka		katsiŋka
FIREWOOD	tsitsi	tsʰitsʰi	tsʰitsʰi	tsitsi		tsitsi
URINATE	tsin		tsʰint	tsint	tsint	tsint

Table 42: **ts* > *tsʰ* / *_i* in Ashéninka and Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
Accipitridae spp.	pagitsa	pakitʰa	pakitʰa	pakitsa	pamakaβiɾiβakitsate	pakitsa
put on garment	kitsagá	kitʰaa	kitʰaa	kitsaa	kitsaa	kitsa
HIGH, tall	tsantsáa	santʰa	hantʰa	tsantsa	katsantsaheɣi	gatsantsani
thread classifier	-tsa	-tʰa	-tʰa	-tsa	-tsa	-tsa
VEIN, artery	sitsa	ʃitʰapi	ʃitʰapaeɾiki	ʃitsa	ʃitsaki	ʃitsa
Heteropsis spp.	topetsa	tapetʰa	tapetʰa	tapetsa	tapetsa	tapetsa
meat, flesh	batsa	βatʰa	watʰa	βatsa		batsa
EGG	gitso	oitʰoki	iitʰoki	itsoki	itsoki	gitsoki
FINISH (vi.)-1	tsoŋ	tʰoŋk	tʰoŋkapaak	tsoŋk		tsoŋka
LEOPARDUS PARDALIS	matsóntsoɾi	matʰontʰoɾi	matʰontʰoɾi	matsontsoɾi		matsontsoɾi
SUCK	tsotegá	tʰo	tʰo	atso		tso
SHARP		tʰojempi	tʰojempi	tsampi	tsampi	tsojampita
BREAST	tsómi		tʰome	tsomi		tsomi

Table 43: **ts* > *tʰ* / *_a, _o* in Ashéninka and Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
THORN	tseí	tʃʰee	kitotʃee	tseeki		tsei
LIP	tséɾá	tʃʰeɾa	tʃeɾa	tseɾa		tseɾa
CUT		tʃʰek	tʃek	tseik		
HOPLIAS SP.	tsiŋoɾi		tʃeŋkoɾi			tseŋkoɾi
HILL, mountain			otʃempi		tsempi	
naked			kaaŋkitʃempoki	kaaɾatsempeki		

Table 44: **ts* > *tʃʰ* / *_e* in Ashéninka, **ts* > *tʃ* / *_e* in Pajonal

4.2.2 **tʃ* > *tʃʰ* in Ashéninka

PK **tʃ* became /tʃʰ/ in Ashéninka but not in Pajonal.

	Nom	Ashé	Paj	Ashá	Kak	Mat
TREE	aɲtʃáto	iɲtʃʰato	iɲtʃato	iɲtʃato	iɲtʃato	iɲtʃato
SOUR	kótʃó	katʃʰo	katʃo	katʃoo	katʃo	katʃo
SWEET	patʃá	potʃʰa	potʃa	potʃa	potʃa	potʃa
ARROW	tʃokopi	tʃʰakopi	tʃekopi	tʃakopi		tʃakopi
YESTERDAY	tʃopi	tʃʰapiŋki	tʃapiŋki		tʃapiŋki	tʃapi
LAGENARIA SICERARIA	patʃaka	patʃʰaka	patʃaka	patʃaka		
GRANDFATHER-1	tʃárine	tʃʰarine	tʃarini	tʃarine		
Carica papaya		mapotʃʰa	mapotʃa	mapotʃa	mapotʃa	
Phaseolus sp.		matʃʰaaki	matʃaaki	matʃaaki	matʃaaki	

Table 45: **tʃ > tʃʰ* in Ashéninka

4.2.3 **t > ts / _i* in Ashéninka and Pajonal

In Ashéninka and Pajonal, PK **t* became /ts/ before *i*, replenishing the phoneme /ts/ which had disappeared following the three-way split described in Section 4.2.1.

	Nom	Ashé	Paj	Ashá	Kak	Mat
BIXA ORELLANA	tsóti	potʰotsi	potʰotsi	potsoti	potʃoti	potsoti
BROTHER-1	íɾenti	iɾentsi	iɾentsi	iɾenti		iɾenti
BASKET	kantiɾi	kantsiɾi	kantsiɾi	kantiɾi		kantiɾi
EXIST	tím	tsim	tsim	tim		tim
FOOT-1	gítí	iitsi	iitsi	iti		giti
IPOMOEA BATATAS	koɾití	koɾitsi	koɾitsi	koɾiti		koɾiti
ORTALIS sp.	maɾáti	maɾatsi	maatsi	maɾati		maɾati
POT	kobiti	koβitsi	kowitsi	koβiti		kobiti
SALT	tibi	tsiβi	tsiwi	tiβi		tibi
CORAGYPS ATRATUS	tisó	tsiso	tsiho	tiso		tiso
FUR	piti	βitsi	witsi	piti		biti
ROTTEN	sítí	ʃitsi	ʃitsi		ʃiti	ʃiti
STAND UP		katsij	katsij	kati	katiɣ	kati

Table 46: **t > ts / _i* in Ashéninka and Pajonal

Table 47 shows several instances in which **t* skipped over /ts/ and went to /tsʰ/ before *i* in Ashéninka. In the words for "DOG" and "Gryllidae spp.", this can be explained as assimilation with a PK **ts* earlier in the word. For the other cases, a possible explanation is that the *t > ts > tsʰ* chain shift before *i* did not occur as two neatly ordered sound changes. Rather, perhaps the **ts > tsʰ* change was still operative when the **t > ts* change arose, allowing some instances of PK **t* to undergo both changes and end up as /tsʰ/.

Curiously, in the last two cognate sets, Ashéninka contains doublets in which one word has /ts/ and another has /tsʰ/. /tsija/ means "FECES" while /tsʰija-aki/ means "ANUS", and /ʃintsipaa/ means "RAFT" while /ʃintsʰipaa/ means "OCHROMA PYRAMIDALE". This may be the result of dialectal variation.

I also found one possible instance of **t > tsʰ / _i* in Pajonal: the word for "Formicidae sp." is /katsʰikoɾi/ while it is /katikoɾi/ in Asháninka and Matsigenka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
NEPHEW-3	tíneɾi	tsʰineɾi	tsineɾi	tinentsi	tineɾi	tineɾi
NEPHEW-1		ɲotsʰi	ɲotsi	nioti		
DRAGONFLY	síkenti	ʃijentsʰi	ʃijentsi	ʃijenti		ʃigenti
DOG	otsíti	otsʰitsʰi	otsʰitsi	otsiti		otsiti
Gryllidae spp.	tsíti	tsʰitsʰi				
FECES	tigá	tsija, tsʰija-aki	tsija	tia	tia	tiga
RAFT	sintipoá	ʃintsipaa, ʃintsʰipaa	ʃintsipaa	ʃintipa		ʃintipoa

Table 47: Sporadic *t > ts^h / _i in Ashéninka

4.3 Diversification of Proto-Kampa *m

In general, PK *m remains as /m/ in all of the daughter varieties.

4.4 Diversification of Proto-Kampa *n

In general, PK *n remains as /n/ in all of the daughter varieties.

4.4.1 /ɲ/ in Kakinte and Matsigenka

/ɲ/ is a marginal phoneme in Kakinte and Matsigenka. This table shows the only two cognate sets in which /ɲ/ in Kakinte or Matsigenka can be put in correspondence with a sound in another variety. In both of these cases, /ɲ/ comes from a palatalized PK *n, but this does not appear to be a regular sound change, as Table 33 shows unpalatalized /n/ in Kakinte /aniani/ and Matsigenka /nia/. It is possible that /maɲo/ and /pooɲa/ were borrowed from a variety of Ashéninka, in which the palatalization is regular.

The other instances of /ɲ/ in Kakinte and Matsigenka have no cognates in the data, so it is impossible to tell whether they result from sound change or borrowing.

mosquito	monio	maɲoo	maɲo	maneo		maɲo
CACICUS SP.	poonia	pooɲa	pooɲa		pooɲa	

Table 48: Sporadic *n > $ɲ$ in Kakinte and Matsigenka

4.5 Diversification of Proto-Kampa *p

In general, PK *p remains as /p/ in all of the daughter varieties.

4.5.1 *p > ∅ / N_ in Nomatsigenga

Proto-Kampa *Np clusters, which were realized as *mp, were generally reduced to /m/ in Nomatsigenga. This was previously noted by Michael et al. (Michael et al. 2010, p. 3). The second table shows several exceptions in which *Np was retained in Nomatsigenga. I cannot find a phonological environment that explains these cases, so my preliminary hypothesis is that they represent borrowings from neighboring varieties of Ashéninka or Asháninka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
EAR	gemitá	kempita	kempita	kempita	tʃoɣempiki	gempita
SAND	omaniki	impaneki	impaneki	impaneki		impaneki
SHOULDER	asémá		ʃempa	sempa		sempa
PULL-1	timá			timpa		
Tabanidae spp.	símokí	ʃimpoki	ʃimpoki	ʃimpoki	ʃimpokiti	ʃimpokiti
ask (question)-1	osámi	osampi	ohampi	sampi		
Gossypium barbadense	omegí	ampee	ampee	ampehi	mampe	ampei
LACHESIS MUTA	kimiɾo	kempiɾo	kempiɾo			kempiɾo
LEFT (HAND)	máte		ampate	nampate		nampate
INSIDE	tsomagi			tsompoi	tsompoɣi	tsompogi
INTESTINES-1	maɾétsá	amporetʰo		amporetsa		poretsa
Loricariidae spp.-1	kimiti	kempitsi		kempiti		kempiti

Table 49: **p* > ∅ / *N_* in Nomatsigenga

	Nom	Ashé	Paj	Ashá	Kak	Mat
PODOCNEMIS UNIFILIS	sempiɾi	ʃempiɾi	ʃempiɾi	sempiɾi		sempiɾi
Banisteriopsis caapi	komaɾampi	kamaɾampi	kamaampi	kamaɾampi		kamaɾampi
Saccharum sp.-2	ompoko					impogo
MEDICINE	hompí			hampi		ampi
HIT, beat	ompos	ompos	ompoh		ompos	
PLAY	mágempí		majempi	majempi		magempi

Table 50: Retentions of **Np* in Nomatsigenga

4.6 Diversification of Proto-Kampa **b*

PK **b* remains as /b/ in Nomatsigenga and Matsigenka. It was lenited to /β/ in Ashéninka, Asháninka, and Kakinte, and it was lenited to /w/ in Pajonal.

	Nom	Ashé	Paj	Ashá	Kak	Mat
Gynerium sagittatum	soboɾo	saβoɾo	hawoo	saβoɾo	saβoɾo	saboɾo
pass (vi.)	ábis	aβis	awih	aβis	aβis	abis
DAUGHTER-IN-LAW	ebá	eβataiɾo	ewatʲeeɾo	eβatajeɾo	iβatʲaɣeo	nebatʲage
WING	ʃíbáŋí	ʃiβaŋki	ʃiwaŋki	ʃiβaŋki	ʃiβaŋki	
durative	-bági	-βai	-wae	-βee		-bage
CURE, HEAL (vt.)		aaβent	aawint	aaβint	ahaβint	ogabint
DECEIVE	matobí		amatawi	amataβi	amataβih	amatabi
POT	kobiti	koβitsi	kowitsi	koβiti		kobiti
capybara	ibeto	iβeto	iweto	iβeto		ibeto
MAKE	obetsik	oβetsʰik	wetsʰik	oβetsik		obetsik

Table 51: **b* > *β* in Ashéninka, Asháninka, and Kakinte, **b* > *w* in Pajonal

4.6.1 Sporadic Loss of **b* in Pajonal and Asháninka

There is evidence of the sporadic sound change **b* > ∅ in Pajonal and Asháninka, and one possible instance in Ashéninka (the cognate set is not very convincing). This is shown in the first three rows of the table. The next three rows show instances in which **b* was lost in Pajonal (causing vowel coalescence) but retained in Asháninka. The last row shows an instance in which **b* was retained as

/w/ in Pajonal but lost in Asháninka. I do not have an explanation for these phenomena, and further investigation is needed.

	Nom	Ashé	Paj	Ashá	Kak	Mat
SHADOW, shade			aampaɾi	ampaɾe		bampaɾoki
illuminate (vt.)-1				tsio		tsibo
illuminate (vt.)-2	obo	oo	oo	oo		
grow (only plants) (vi.)			ʃook	ʃiβo		ʃibok
imitate (animal cry, person)			aako	oβako		
indirect appl. + benefactive appl.			aka-ant	ako-βent		
drop (vt.)	obaɾíg		owaɾij	oaɾi		

Table 52: Sporadic losses of **b* in Pajonal and Asháninka

4.7 Diversification of Proto-Kampa **t*

In general, PK **t* remains as /t/ in all of the daughter varieties.

4.7.1 /tʲ/ in Matsigenka

/tʲ/ is a marginal phoneme in Matsigenka, only appearing ten times in my data set. Table 53 shows the number of occurrences of /tʲ/ and /ti/ before the vowels /a/, /e/, and /o/ in Matsigenka. Table 54 shows the only instances in which /tʲ/ in Matsigenka can be put in correspondence with a sound in another variety; in those cases, it comes from palatalization of PK **t* before the sequence *io*. Thus, the simplest explanation for the data is that PK **t* uniformly palatalized before *iV* sequences in Matsigenka, producing the phoneme /tʲ/. The one instance of /tia/ is in the word /motia/ meaning "STOMACH" (see Table 36). We see that the PK form is **motiha*, and the presence of /tia/ instead of /tʲa/ in Matsigenka can be explained by positing that palatalization of **t* preceded loss of **h* in Matsigenka.

	_a	_e	_o
tʲ	2	0	8
ti	1	0	0

Table 53: Occurrences of /tʲ/ and /ti/ before /a/, /e/, and /o/ in Matsigenka

	Nom	Ashé	Paj	Ashá	Kak	Mat
DUCK		paɲtʃo	pantʲo	pantio		pantʲo
SMALL-2	tiómihání					tʲomiani
BUTTOCKS	tiogi			tioki		tʲoki
SISTER, WOMAN'S		eeɲtʃo	eentʲo			entʲone

Table 54: **t* > *tʲ / _io* in Matsigenka

4.7.2 /tʲ/ in Kakinte

The phoneme /tʲ/ is considerably more common in Kakinte than in Matsigenka, but due to the considerable divergence of the Kakinte lexicon, there are only seven instances in the data where /tʲ/ in Kakinte can be put in correspondence with a sound in another variety. These are shown in the table below. In the first five rows, /tʲ/ in Kakinte corresponds to PK **t* and there is no clear cause

for palatalization. In /kitʲoŋka/, the /tʲ/ appears to come from a PK **ti* sequence. In /iβatʲaɣeo/, the /tʲ/ likely comes from affective palatalization, which often affects consonants in familial terms throughout the Kampa family (Payne 1981, p. 50).

A possible explanation for Kakinte's development of /tʲ/ and its lexical divergence is that Kakinte speakers have engaged in esoterogeny, defined by Dimmendaal as "a phenomenon whereby speakers add linguistic innovations that increase the complexity of the language in order to highlight their distinctiveness from neighboring groups", in this case Asháninka speakers (Dimmendaal 2015, p. 64). More research on the social, political, and linguistic history of the Kakinte community is needed to better understand this phenomenon.

	Nom	Ashé	Paj	Ashá	Kak	Mat
NAIL, CLAW	ʃátá	ʃetaki	ʃeta	setaki	ʃatʲaki	ʃata
upriver, upstream	toóŋkí	katoŋko	katoŋko	katoŋko	katʲoŋko	katoŋko
shin			tawaato	taβaato	tʲaβaato	tabatokii
CATERPILLAR		kotaa			kotʲaʃi	
PLANT (V.)			owint		oβintʲak	
RED		kitʃoŋka	kitʲoŋka		kitʲoŋka	
DAUGHTER-IN-LAW	ebá	eβataiɾo	ewatʲeeɾo	eβatajeɾo	iβatʲaɣeo	nebatʲage

Table 55: /tʲ/ in Kakinte

4.8 Diversification of Proto-Kampa **k*

In general, PK **k* remains as /k/ in all of the daughter varieties.

4.8.1 **k* > ∅ / *N_* in Nomatsigenga

Proto-Kampa **Nk* clusters, which were realized as **ŋk*, were generally reduced to /ŋ/ in Nomatsigenga. This sound change created the phoneme /ŋ/, which among Kampa varieties is unique to Nomatsigenga. As noted by Michael, this sound change is parallel to the **Np* > *m* change discussed in Section 4.5.1, but an analogous change did not affect **Nt* sequences (Michael 2011).

The second table shows several exceptions in which **Nk* was retained in Nomatsigenga. These retentions demonstrate that /ŋ/ must be analyzed as a phoneme in Nomatisgenka and not as underlying /Nk/. As with the retentions of **Np*, I cannot find a phonological environment that explains these cases, so my preliminary hypothesis is that they represent borrowings from neighboring varieties of Ashéninka or Asháninka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
air classifier	-heŋa	-iŋka	-eŋka	-heŋka	-heŋka	-eŋka
BLOW	tasóŋ	tasoŋk	tahoŋk	tasoŋk	tasoŋk	tasoŋk
BONE	toŋi	toŋki	toŋki	toŋki	toŋki	toŋki
CORN	síŋí	ʃiŋki	ʃiŋki	ʃiŋki	ʃiŋki	ʃiŋki
CLIMB-1	toŋá	toŋkaɣ	toŋka	toŋka	toŋkoɣ	toŋko
FATHER-IN-LAW	koŋiɾi	koŋkiɾi	koŋkiɾi	koŋkiɾi	kooŋkini	koŋkiɾi
HOUSE	poŋo	paŋko	paŋko	paŋko	paŋko	paŋko
PUSH	otatsiŋ	tatsʰiŋk	otatsʰiŋk	tatsiŋk	tatsiŋk	tatsiŋk
RAIN	aŋani	iŋkane	iŋkani	iŋkani	iŋkani	iŋkani
REMEMBER	keŋa	kiŋkiʃiɾʲ	keŋkitʰaʃiɾʲ	keŋkeʃiɾe	keŋkeh	keŋki
VOMIT	komaɾaŋá	kamaɾaŋk	kamaaŋk	kamaɾaŋk	kamaɾaŋk	kamaɾaŋk

Table 56: **k* > ∅ / *N_* in Nomatsigenga

	Nom	Ashé	Paj	Ashá	Kak	Mat
Saccharum sp.-1	soŋko	saŋko	haŋko	saŋko	saŋko	ʃaŋko
upriver, upstream	toóŋkɨ́	katoŋko	katoŋko	katoŋko	katʲoŋko	katoŋko
FLUTE-1	soŋkaɾi	soŋkaɾi	hoŋkamento		soŋkaɾi	soŋka
TROCHILIDAE SPP.	tsoŋkiɾi	tʰoŋkiɾi	tʰoŋkiɾi	tʃoŋkihi		tsoŋkiɾi
BEAD	niŋketsiki		neŋketsʰiki	neŋke		neŋketsiki
CHIEF	piŋkatsáɾi	piŋkatʰáɾi	piŋkatʰáɾi	piŋkatsáɾi		
HEART	asaŋkane	asaŋkane	ahaŋkane	asaŋkane		
FOOT-2	boŋkítí	poŋkitsi			ʃiβoŋkiti	
DANCE	mániŋke			maniŋke		

Table 57: Retentions of **Nk* in Nomatsigenga

4.9 Diversification of Proto-Kampa **g*

In general, PK **g* remains as */g/* in Nomatsigenga and Matsigenka. The development of **g* in Ashéninka, Pajonal, Asháninka, and Kakinte is discussed below.

4.9.1 Lenition and Loss of **g* in Asháninka

In Asháninka, PK **g* went to */j/* before *e* and was lost elsewhere, as described by Michael et al. (Michael et al. 2010, p. 2). The table below shows examples of **g > j / _e*. Many examples of **g > ∅* can be found in Section 4.9.2.

	Nom	Ashé	Paj	Ashá	Kak	Mat
DRAGONFLY	síkenti	ʃijentsʰi	ʃijentsi	ʃijenti		ʃigenti
PLAY	mágempí		majempi	majempi		magempi
KNEE	geɾíto	jeɾeto	jeɾito	jeɾito		geɾeto
KNEEL				tijeɾo		tigeɾo
BROTHER-2	igé	ije		ijenti	iɣenti	ige
ISLAND	obógesi			oβojeha		obogea
Saimiri sp.	tsigiɾi	tsʰijeɾi	tsʰijeɾi	tsijeɾi	tsieɾi	tsigeɾi

Table 58: **g > j / _e* in Asháninka

4.9.2 Lenition and Loss of **g* in Ashéninka and Pajonal

The following table shows the most common reflex of PK **g* in Ashéninka and Pajonal in intervocalic environments, with the row indicating the preceding vowel and the column indicating the following vowel. An empty cell signifies that no cognate set with that environment was found in my data. β/w signifies /β/ in Ashéninka and /w/ in Pajonal, and /w means no data in Ashéninka and /w/ in Pajonal. It would be unwise to give too much weight to this table, as exceptions exist and some cells are based on just one cognate set.

The end of a verb root behaves like a *_a* environment because I extracted Ashéninka and Pajonal verb roots from the nominalized form, which has the suffix /-aantsʰi/ from PK **-agantsi*. In many cases, these verb roots have other allomorphs in other environments, so the form presented here is not necessarily the correct underlying form. Time constraints prevented me from studying this allomorphy, but a detailed analysis could shed more light on the development of **g* in Ashéninka and Pajonal.

	_a	*_o*	*_ui*	*_e*	*_i*
*a_*	∅	β/w	∅	j	∅
*o_*	β/w	∅	∅	j	∅
*ui_*					
*e_*	j	/w	∅		j
*i_*	j	j		j	∅

Table 59: Development of intervocalic **g* in Ashéninka and Pajonal

Previously, Michael et al. argued that in Ashéninka, **g* was lost following *a* or between identical vowels, and **g* became /j/ elsewhere (Michael et al. 2010, p. 2). This description fails to account for the loss of **g* in the *o_ui*, *o_i*, and *e_i* environments and does not mention **g* > *β* or **g* > *w*.

Instead, I propose the following analysis. In Ashéninka and Pajonal, PK **g* lenited in all intervocalic environments, and it was able to disappear completely if the resulting vowel sequence was acceptable. Specifically, **aga*, **ogo*, and **igi* became the long vowels /aa/, /oo/, and /ii/. Assuming that the **ui* > *i* merger had already occurred, **agi* and **agui* sequences became the diphthong /ai/ (which later became /ae/ in Pajonal), while **ogi* and **ogui* sequences became the diphthong /oi/ (which later became /oe/ in Pajonal). In other environments, loss of **g* would lead to an impermissible vowel sequence, so it developed into a glide. Specifically, **g* became /j/ if adjacent to a front vowel. Otherwise, **g* became /w/ in Pajonal, and in Ashéninka **g* became the glide [β̞] in the *a_o* environment and [w] in the *o_a* environment (these are allophones of /β/ (Payne 1980, p. 10)).

This analysis does not explain all of the data. For example, it does not predict **ego* > *ewo* in the Pajonal word /kewo/ (see Table 61) or **egui* > *ee* in the Ashéninka and Pajonal words /ʃeeto/ (see Table 62). I am also purposely vague in invoking the notion of acceptable vowel sequences, as a proper analysis would require a thorough investigation of the history of vowels in Ashéninka and Pajonal.

Interestingly, this analysis suggests that **g*-loss, or at least its final stages, may have occurred after the palatalization described in 4.1 eliminated **ea*, **ia*, and **io* sequences. For if such sequences still existed and **g* tended to delete completely wherever possible, we would expect **ega* > *ea* instead of **ega* > *eja* for example.

The following tables demonstrate the development of intervocalic **g* in Ashéninka and Pajonal. The first table shows examples of **g* > *j* in the environments indicated by Table 59, as well as one instance of **agi* > *aji*. The second table shows that **g* becomes /β/ in Ashéninka and /w/ in Pajonal in *a_o* and *o_a* environments, as well as one possible instance of **ego* > *ewo* in Pajonal. The third table shows **g* > ∅ in other intervocalic environments.

	Nom	Ashé	Paj	Ashá	Kak	Mat
FLOWER	otega	teja	teja	teaki	otʃotenteɣa	tega
OENOCARPUS BATAUA	sega	ʃeja	ʃeja	ʃaa	ʃeɣa	sega
RETURN (vi.)	pig	pij	pij	pi	piɣ	pig
swallow (v.)	nig	nij	nij	ni	niɣ	nig
DIG	kig	kij	kij	ki		kig
sick, ill-1	mantsigá	mants^{h}ija	mants^{h}ija	mantsija		mantsiga
MIDDLE		nijaŋki	nijaŋki	niaŋki	niɣaŋki	nigaŋki
STAND UP		katsij	katsij	kati	katiɣ	katig
PSEUDOPLATYSTOMA SP.	igotsi	ijots^{h}i	ijots^{h}i	iotsi		igotsi
DRAGONFLY	síkenti	ʃijents^{h}i	ʃijentsi	ʃijenti		ʃigenti
Saimiri sp.	tsigiɾi	ts^{h}ijeɾi	ts^{h}ijeɾi	tsijeɾi	tsieɾi	tsigeɾi
Sciuridae spp.		mejiɾi	mejiɾi	meiɾi	meiɾi	megiɾi
PLAY	mágempí		majempi	majempi		magempi
RAINBOW	óge	oje	oje			joge
MOTHER-IN-LAW-1	ágiɾo	ajiɾo	ajiɾo	aiɾo	aiɾo	pagiɾo

Table 60: Intervocalic **g* > *j* in Ashéninka and Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
DIOSCOREA SP.	mágona	maβona	mawona	maona	maona	magona
CLOTH	magóɾi		mawots^{h}i			pamagotiɾo
KILL	ogámag	oβamaɣ	owamaɣ			ogamag
squeeze (vt.)		oɲtʃ^{h}ooβ	awiʃoow	aβitso		obitʃog
EAT-1		oβ	ow			og
light a fire (vt.)	ogamá		owaama			ogima
ask (question)-2			kowako		koako	kogakotagant
SAD	ogasɨɾe		owaʃiɾe	oaʃiɾe		
HAMMOCK			kewo			gegontamento

Table 61: **g* > *β / a_o, o_a* in Ashéninka, **g* > *w / a_o, o_a, e_o* in Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
ALL	ómagaɾo	maaɾoni	maawaeni	maaɾoni	maasano	magaɾoni
demon, devil	kámagaɾi	kamaaɾi	kamaaɾi	kamaaɾi	kamaaɾi	kamagaɾini
INFLUENZA, COLD	kámagantsi	kamaants^{h}i	kamants^{h}i	kamaantsi		
PLANT SP.	pótogo	potoo	potoo	potoo	potoɣo	potogo
INTESTINES-2	ségóto	ʃeeto	ʃeeto	ʃeito	seito	segɯito
NAVEL	tʃomagíto	moito	maeto	moiɾo	moito	mogɯito
TIE-1	ogɨʃó	ooso	ooho	oiso	oiʃo	ogɯisot
CLIMB-2		atai	atee	atai	atai	atagɯi
SIMULIUM SPP.	tsigito	ts^{h}iito		tsiito	tsiito	tsigito
DAY	katagítéɾi	kitaite		kitaiteɾi		kɯitagiteɾi
return	ogipig		oepij	oipi		ogipig

Table 62: Intervocalic **g* > ∅ in Ashéninka and Pajonal

The following table shows the development of **g* in Ashéninka and Pajonal at the beginning of noun roots. The first three rows show that Ashéninka and Pajonal have /ii/ as the reflex of root-initial **gi*. This could be the result of **g*-loss and compensatory lengthening, or it could indicate that the third

person masculine possessed forms, which begin with /ii/ from PK **i-gi*, were reanalyzed as the root. The last row shows /je/ as the reflex of root-initial **ge* in Ashéninka and Pajonal. Again, this could result from a regular sound change, or it could arise from leveling as **g* became /j/ intervocalically in inflected forms, e.g. **no-gerito* > *no-jerito* in Pajonal.

	Nom	Ashé	Paj	Ashá	Kak	Mat
HEAD	gíto	iito	iito	iito	ʃokoito	gito
FOOT-1	gítí	iitsi	iitsi	iti		giti
HAIR, HEAD	gisí		iiʃi	iʃi		giʃi
KNEE	geríto	jeɾeto	jeɾito	jeɾito		geɾeto

Table 63: Development of root-initial **g* in Ashéninka and Pajonal

There is one further complication. PK verb roots ending in **ag* were retained as /aɣ/ in Ashéninka and Pajonal in the nominalized form, from which roots were extracted. The PK nominalizing suffix is **-agantsi*, creating an **ag-aga* environment for these roots. As the **aga* sequence in the suffix became /aa/, the **g* in the root could not delete because it would create the impermissible trimoraic /aaa/. These instances of **g* had no phonetic motivation to become /j/ as they were not adjacent to a front vowel, and they had no phonetic motivation to become /β/ or /w/ as they were not adjacent to *o*. Thus, they were retained as /ɣ/.

In addition, in a number of cases Ashéninka and Pajonal appear to have innovated a root final /ɣ/ when it was not present in Proto-Kampa. (Alternatively, it is possible that these cases show loss of **g* in Nomatsigenga and Matsigenka.) These root-final /aɣ/ sequences account for all instances of /ɣ/ in my data for Ashéninka and the majority of instances in Pajonal.

	Nom	Ashé	Paj	Ashá	Kak	Mat
TAKE (vt.)	ág	aɣ	aɣ	aa	aɣ	ag
KILL	ogámag	oβamaɣ	owamaɣ			ogamag
SLEEP (V.)		maɣ	maɣ	ma	maɣ	mag
CRY (v.)	iɾák	iɾaɣ	iɾaɣ	iɾa	iɾaɣ	iɾag
BURN	ta	taɣ	taɣ	ta	taɣ	tag

Table 64: Retentions of root-final **ag* in Ashéninka and Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
carry		t^{h}omaɣ	t^{h}omaɣ			tsoma
LIFT (vt.)-1	tiná		tsinaɣ			
JUMP-2		mitaɣ		mita		mita
BREAK (vt.)-2	osataá	sataɣ		sata		
TEAR (vt.)	saɾaá	saɾaɣ				saɾa
WAIT	ogia	oj	ojaɣ	oja	oɣih	ogi

Table 65: Apparent innovations of root-final /ɣ/ in Ashéninka and Pajonal

4.9.3 Development of **g* in Kakinte

PK **g* either becomes /ɣ/ or is lost in Kakinte. The conditioning environment for the loss is unclear to me. Unlike in Ashéninka, Pajonal, and Asháninka, it does not appear to be largely characterized by the surrounding vowels. The following tables show instances of **g* being retained and lost in Kakinte in the same intervocalic environments. It is possible that some other factor explains the patterns in

the following tables, or that /ɣ/-loss is an ongoing process in Kakinte that is spreading through the lexicon.

	Nom	Ashé	Paj	Ashá	Kak	Mat
MOUTH	bagante	paante	paante	βante	βaɣante	bagante
BE ABLE		aaβij		aβe	aɣaβeh	agabe
FINISH (vi.)-2			aat	aat	aɣa	agat
CLEAN				kaharaite	kaharaɣite	
CHEST	neí			nehi	neɣi	negi
MIDDLE		nijaŋki	nijaŋki	niaŋki	niɣaŋki	nigaŋki
BROTHER-2	igé	ije		ijenti	iɣenti	ige
LIFT (vt.)-2	ogáinok				oɣahenok	ogaenok

Table 66: **g* > *ɣ / a_a, a_i, e_i, i_a, i_e, o_a* in Kakinte

	Nom	Ashé	Paj	Ashá	Kak	Mat
ALL	ómagaro	maaroni	maawaeni	maaroni	maasano	magaroni
demon, devil	kámagari	kamaari	kamaari	kamaari	kamaari	kamagarini
MOTHER-IN-LAW-1	ágiro	ajiro	ajiro	airo	airo	pagiro
Sciuridae spp.		mejiri	mejiri	meiri	meiri	megiri
FECES	tigá	tsija	tsija	tia	tia	tiga
Saimiri sp.	tsigiri	ts^{h}ijeri	ts^{h}ijeri	tsijeri	tsieri	tsigeri
ask (question)-2			kowako		koako	kogakotagant

Table 67: **g* > ∅ */ a_a, a_i, e_i, i_a, i_e, o_a* in Kakinte

4.9.4 Sporadic Loss of **g* in Nomatsigenga

There are several instances of PK **g* being lost in Nomatsigenga. I see no conditioning environment that unifies these cases, so the best explanation I can give is that this is the result of borrowing from Asháninka or a variety of Ashéninka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
BURN	ta	taɣ	taɣ	ta	taɣ	tag
LOSE (vt.)	pí	pej	pej	pe	pe	peg
SHAMAN	seripiári		ʃeripijari			seripigara
CHEST	neí			nehi	neɣi	negi
BACK (n.)	tapíi	taapii	tapii	tapii		tapigi
Lagothrix sp.	komaínárí					komaginaro

Table 68: Sporadic losses of **g* in Nomatsigenga

4.10 Diversification of Proto-Kampa **ts*

In general, PK **ts* remains as /ts/ in Nomatsigenga, Asháninka, Kakinte, and Matsigenka. The development of **ts* in Ashéninka and Pajonal is discussed in Section 4.2.1.

4.11 Diversification of Proto-Kampa **tʃ*

In general, PK **tʃ* remains as /tʃ/ in all of the daughter varieties, except in Ashéninka in which it becomes /tʃ^{h}/. See Table 45 for examples.

4.12 Diversification of Proto-Kampa **s*

In a non-palatalizing environment, PK **s* remains as /s/ in all of the daughter varieties, except in Pajonal in which it becomes /h/, merging with the already existing phoneme /h/. According to Pedrós, this **s* > *h* change is a distinguishing feature of the Pajonal and Ucayali varieties of Ashéninka (Pedrós 2018, p. 11).

	Nom	Ashé	Paj	Ashá	Kak	Mat
BLOW	tasóŋ	tasoŋk	tahoŋk	tasoŋk	tasoŋk	tasoŋk
Gynerium sagittatum	soboɾo	saβoɾo	hawoo	saβoɾo	saβoɾo	saboɾo
pass (vi.)	ábis	aβis	awih	aβis	aβis	abis
SWEAT	masobí	masaβi	mahawi	masaβi		masobi
CORAGYPS ATRATUS	tisó	tsiso	tsiho	tiso		tiso
GRANDFATHER-2	sárí	saɾi	haɾi	saɾi		pisaɾi
KIDNEY	sóŋíɾékí	soŋki	hoŋki	soŋkepeo		soŋkipegoki
WASP	sani	sanii	hani	sani		sani
FLUTE-1	soŋkaɾi	soŋkaɾi	hoŋkamento		soŋkaɾi	soŋka

Table 69: **s* > *h* in Pajonal

4.12.1 **s* > *ʃ / _i* Except in Nomatsigenga

The following table shows the development of PK **si* and **sui* in the daughter varieties. Ashéninka, Pajonal, Asháninka, and Kakinte show **sui* > *ʃi*. The only reasonable sequence of sound changes for this to occur is **sui* > *si* > *ʃi*. Because these varieties also have **si* > *ʃi*, the most parsimonious explanation is that **s* > *ʃ / _i* happened once in the history of each of these varieties, and that it was fed by **ui* > *i*. If we accept this, then the fact that **ui* > *i* did not happen in Matsigenka means that **s* > *ʃ / _i* must have occurred independently in Matsigenka, which is quite plausible because it is a common sound change cross-linguistically.

PK	Nom	Ashé	Paj	Ashá	Kak	Mat
**si*	si	ʃi	ʃi	ʃi	ʃi	ʃi
**sui*	sɨ	ʃi	ʃi	ʃi	ʃi	sɯi

Table 70: Development of PK **si* and **sui*

This contradicts Michael's suggestion that **s* > *ʃ / _i* is a shared innovation among Matsigenka, Nanti, Kakinte, Asháninka, and Ashéninka (Michael 2011). If **s* > *ʃ / _i* were indeed a shared innovation among non-Nomatsigenga Kampa varieties, we would need to say that the palatalization happened again in Ashéninka, Pajonal, Asháninka, and Kakinte after **ui* > *i* occurred, which is unnecessarily complicated.

The following tables show that **s* > *ʃ / _i* occurred in every variety in my sample besides Nomatsigenga, and that out of those varieties, it was fed by **ui* > *i* in every variety besides Matsigenka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
RUN	síg	ʃij	ʃij	ʃi	ʃiɣ	ʃig
ROAST	tasi	taʃi	taʃi	taʃi	taʃi	taʃi
delicious	basíní	poʃini	poʃini	poʃini	poʃini	poʃini
DEFECATE	sí	ʃi	ʃi	ʃi	ʃi	ʃi
PULL-2	nósik	noʃik	noʃik	noʃik	noʃik	noʃik
TIE-2	sitíg	ʃiɾik	ʃiɾik	ʃitik	ʃitik	ʃitik
Tabanidae spp.	símokí	ʃimpoki	ʃimpoki	ʃimpoki	ʃimpokiti	ʃimpokiti
VEIN, artery	sitsa	ʃit^{h}api	ʃit^{h}apaeɾiki	ʃitsa	ʃitsaki	ʃitsa
LEAF	osí	ʃi	ʃi	ʃi	ʃi	ʃi
CORN	síŋí	ʃiŋki	ʃiŋki	ʃiŋki	ʃiŋki	ʃiŋki

Table 71: **s* > *ʃ / _i* except in Nomatsigenga

	Nom	Ashé	Paj	Ashá	Kak	Mat
LAUGH	síɾon	ʃiɾont	ʃiɾonta	ʃiɾont	ʃiɾont	ɯiɾont
MAN	siɾáɾi	ʃiɾampaɾi	ʃiɾampaɾi	ʃiɾampaɾi	ʃiɾaβaɾi	ɯiɾaɾi
THINK	síɾe	kiŋkiʃiɾ^{j}	keŋkit^{h}aʃiɾ^{j}	keŋkeʃiɾe		ɯiɾe
WHISTLE (V.)	síba			ʃiβa	ʃiβa	ɯibat
SAD	ogasíɾe		owaʃiɾe	oaʃiɾe		

Table 72: **ɯi* > *i* fed **s* > *ʃ / _i* in Ashéninka, Pajonal, Asháninka, and Kakinte

4.12.2 **s* > *ʃ / _e* in Ashéninka, Pajonal, and Asháninka

In Ashéninka, Pajonal, and Asháninka, PK **s* additionally palatalized to /ʃ/ before *e*. As shown in the first four rows of the table below, Asháninka has several pairs of forms with /s/ and /ʃ/. In the cognate sets for "HUNGRY", "CRAB", and "PODOCNEMIS UNIFILIS", we may assume that this is due to dialectal variation. In the word for shoulder, Asháninka has apparently preserved /s/ in the unpossessed form /sempa-tsi/ but has /ʃ/ in the possessed form /no-ʃempa/.

Out of the six reliable cognate sets in my data showing a reflex of PK **se* in Kakinte, four have /s/ and two have /ʃ/. This could reflect sociolinguistic or dialectal variation, borrowing from varieties of Ashéninka or Asháninka, or the possibility that the sound change is still making its way through the lexicon in Kakinte. This data contradicts Michael's claim that **s* > *ʃ / _e* happened uniformly in Kakinte (Michael 2011). However, it is clear that he was working with different lexical data, as he has /ʃeɾi/ as the Kakinte word for "Nicotiana tabacum". Once again, this discrepancy could result from sociolinguistic or dialectal variation within Kakinte.

	Nom	Ashé	Paj	Ashá	Kak	Mat
HUNGRY	taség	taʃ	taʃe	tase, taʃe	taseɣ	taseg
CRAB	ósiɾo	oʃeɾo	oʃeɾo	oseɾo, oʃeɾo		oseɾo
PODOCNEMIS UNIFILIS	sempiɾi	ʃempiɾi	ʃempiɾi	sempiɾi, ʃempiɾi		sempiɾi
SHOULDER	sémá		ʃempa	sempa-tsi, no-ʃempa		sempa
Nicotiana tabacum	seɾi	ʃeɾi	ʃeɾi	ʃeɾi	seeɾi	seɾi
INTESTINES-2	ségóto	ʃeeto	ʃeeto	ʃeito	seito	segɯito
ADAM'S APPLE		ʃeno	ʃeno	ʃeno	seno	
PANTHERA ONCA	sekáɾi	kaʃekaɾi	kaʃeekaɾi			
OENOCARPUS BATAUA	sega	ʃeja	ʃeja	ʃaa	ʃeɣa	sega
EAT-2					ʃeka	seka

Table 73: **s* > *ʃ / _e* in Ashéninka, Pajonal, and Asháninka

4.13 Diversification of Proto-Kampa *ʃ

In general, PK *ʃ remains as /ʃ/ in all of the daughter varieties.

	Nom	Ashé	Paj	Ashá	Kak	Mat
DASYPROCTA SP.	ʃaɾo	ʃaɾo	ʃawo			ʃaɾoni
granddaughter-1	ʃóɾó	ʃaβo		ʃaβo	ʃao	
	ʃoŋ		ʃoŋk	ʃoŋk		ʃoŋk
FLUTE-2	ʃobiɾe	ʃoβiɾe	ʃowiɾe	ʃioβiɾe		
HIP	ʃóɾíta			ʃoɾita	ʃoɾita	ʃoɾita
INIA spp.		koʃoʃiko	koʃoʃiko	koʃoʃiko		
SPIT (v.)-1				kiʃok	kiʃoh	kɯiʃok

Table 74: Cognate sets showing *ʃ

4.14 Diversification of Proto-Kampa *h

Intervocalically, PK *h was generally retained in Nomatsigenga, Asháninka, and Kakinte, and lost in Ashéninka, Pajonal, and Matsigenka.

	Nom	Ashé	Paj	Ashá	Kak	Mat
NAME (n.)	pahíɾó	βaiɾo	waeɾo	βahiɾo	pahiɾo	baiɾo
WATER	nihá	ɲaa	ɲaa	niha	niha	nia
garbage	kahárá	kaaɾaʃi	kaaɣaʃi	kahaɾa	kahaɾaʃiteki	kaaɾa
ITCH	kahíní	kaini	kaeni	kaheni		kaeni
TERMITE	kahíɾó	kaiɾo	kaeɾo	kahiɾo		kaiɾo
grab, grasp		oiɾik	oeɾik	ahiɾik	ahiɾik	aiɾik
SILENT, quiet	mahíɾé		maeɾi	maaheɾe		maiɾe
APPEAR		koɲaa	koɲaa		koniha	konea

Table 75: Intervocalic *h > ∅ in Ashéninka, Pajonal, and Matsigenka

In the following pages, I will discuss the development of *h in each variety, ordered from greatest to least retention. I would like to preface this discussion by noting that the synchronic process of optional /h/-deletion has been documented in Kakinte (Swift 1988, p. 132), Nomatsigenga (Lawrence 2013, p. 20), and Nanti (Michael 2008, p. 245). It would be not unreasonable to assume that a similar process occurs in the other Kampa varieties. Whether or not this /h/-deletion occurs is influenced by many factors, including the prosodic environment and speech style, as /h/ is more likely to be retained in emphatic speech (Michael 2008, p. 246). Another relevant fact is that allomorphy between /h/ and ∅ is quite common, especially in verb roots (e.g. /kaíní/-/kahíní/ in Nomatsigenga (Shaver 1996, p. 100) and /moi/-/moih/ in Asháninka (Kindberg 1980, p. 197)). To my knowledge, the causes of this allomorphy and its relation to optional /h/-deletion are not well understood.

One consequence of these facts is that for each instance of apparent *h-loss in my data, we cannot be sure that there is not a synchronic underlying /h/ that has failed to surface due to optional /h/-deletion or some morphophonologcal process. This lack of clarity in the synchronic picture makes it very difficult to formulate a convincing diachronic picture. Thus, I will simply present the relevant data with no presumption that all of the patterns can be explained by historical sound changes.

4.14.1 Development of *h in Kakinte

Out of the varieties in my data set, Kakinte is the most conservative in preserving PK *h. I found only two examples in which *h was apparently lost in Kakinte, and neither cognate set is particularly

convincing. The /-a/ in Ashéninka /konaa/ and the /-ha/ in Asháninka /konaha/ might be extra morphology, namely the fluid classifier. The cognate set for "BROTHER, WOMAN's" only spans two varieties.

Of course, it is possible that some instances of PK **h* were lost in Kakinte as well as in all other varieties in my data set, but I have no way of identifying these cases. Swift describes optional elision of intervocalic /h/ in Kakinte, which would for example cause /arehe/ to surface as [aree] (Swift 1988, p. 132). However, there appears to be no regular sound change involving loss of **h* in the underlying form.

	Nom	Ashé	Paj	Ashá	Kak	Mat
fish (with fish poison)		konaa	kona	konaha	kona	
BROTHER, WOMAN'S				haihi	haai	

Table 76: Losses of **h* in Kakinte

4.14.2 Development of **h* in Asháninka

PK **h* is generally preserved in Asháninka. As shown in the tables below, most instances of lost **h* are at the end of verb roots. However, this is not a steadfast generalization, as root-final **h* is occasionally preserved and non-root-final **h* is occasionally lost. The second table simply lists instances of root-final /h/ in Asháninka; I do not claim that all of the forms should be reconstructed with root-final **h* in Proto-Kampa.

	Nom	Ashé	Paj	Ashá	Kak	Mat
ENTER	k	kʲ	kʲ	kia	kih	ki
REMEMBER	keŋa	kiŋkiʃiɾʲ	keŋkitʰaʃiɾʲ	keŋkeʃire	keŋkeh	keŋki
SEE	nía	ɲ	ɲ	ne	neh	ne
WAIT	ogia	oj	ojaɣ	oja	oɣih	ogi
cross river (v.)	montia	moɲtʃ	montʲaɣ	monte	monteh	monte
lie (on surface)	naɾiá	noɾʲ		naɾe	noɾih	noɾi
break in half				kaβiɾe	kaβiɾeh	
DECEIVE	matobí		amatawi	amataβi	amataβih	amatabi
BREAK (vt.)-1	kará			kaɾa	kaɾah	
BE ABLE		aaβij		aβe	aɣaβeh	agabe

Table 77: Losses of root-final **h* in Asháninka

	Nom	Ashé	Paj	Ashá	Kak	Mat
carry on back	kía	kʲ	kʲaatʰa	kih		ki
split	tsiɾá			tsiɾah	tsiɾah	tsiɾa
BURY	tía		tʲaaki	tih	tihaβio	
BREAK (vt.)-3				βatih	oβateh	obatɯi
BOIL (v.)		moj	moj	moih	moiha	
WEAVE			ontʲaɣ	tih	ontih	
LEAVE			hokʲ	sokih		

Table 78: Instances of root-final /h/ in Asháninka

	Nom	Ashé	Paj	Ashá	Kak	Mat
SMELL (vt.)	kémiheŋá	kemaiŋka	kemaeŋka	kemeeŋka		kemaeŋka
SWIM	maá	amaa	amaa	amaa	amaha	amaa
ARRIVE	aɾe	aɾee	aɾee	aɾee	aɾehe	
sick, ill-2				hokii	ohokihi	
BATHE (vi.)	kaá	kaa	kaawoʃi	ka	kaha	kaa
WALK	anɨí	anii		anii	anihi	anɯii

Table 79: Non-root-final losses of **h* in Asháninka

4.14.3 Development of **h* in Nomatsigenga

PK **h* is generally preserved in Nomatsigenga, but lost more often than in Asháninka. An interesting generalization is that wherever **h* is lost in Asháninka, it is also lost in Nomatsigenga (see Tables 77 and 79). This might suggest the possibility that **h*-loss in Nomatsigenga occurred in two stages: the first in a common ancestor of Nomatsigenga and Asháninka, and the second after they diverged. However, this hypothesis is unlikely as Nomatsigenga and Asháninka do not share a common ancestor after Proto-Kampa according to previous work on subgrouping (see Section 1.4.2).

The following table shows instances in which **h* was lost in Nomatsigenga but retained in Asháninka. These cases include all other instances of root-final **h*, as shown in the first three rows, as well has various intervocalic instances. There is no clear environment that distinguishes these intervocalic losses from the intervocalic retentions shown in Table 75.

	Nom	Ashé	Paj	Ashá	Kak	Mat
carry on back	kía	kʲ	kʲaatʰa	kih		ki
split	tsiɾá			tsiɾah	tsiɾah	tsiɾa
BURY	tía		tʲaaki	tih	tihaβio	
SHOUT, yell	kaím	kaim	kaem	kahem	kahem	kaem
chew	noo	naa	naa	noha	noha	noa
DRY (vi.)-1	piɾiá	piɾʲaa	piɾʲaa	piɾihatahi	piɾih	piɾia
SHOW (V.)	oniág	oɲaak	oɲaaɣ	oniha	onihaɣ	
LAKE	aɲaáɾé	iŋkaaɾe	iŋkaaɾe	iŋkahaɾe		iŋkaaɾe
PARAPONERA CLAVATA	manií	manii	manii	manihi		manii
TOOTH	aí	aiki	aeki	ahi		ai
WIDOW	ogámaímentaga			kamahimentaɾo		
LUNGS	boŋaɾe	βijoŋkaɾe	wʲoŋkaɾe	βihoŋkaɾi		βɯioŋkaɾe

Table 80: Instances of **h*-loss in Nomatsigenga and retention in Asháninka

4.14.4 Development of **h* in Ashéninka and Pajonal

In Ashéninka and Pajonal, PK **h* was lost intervocalically. Because **h* does not occur in syllable codas or in consonant clusters, in practice this means that it was retained only at the beginning of roots. Root-initial retentions of **h* are shown in the table below. The last three rows show instances in which **h* was kept in one of Ashéninka or Pajonal and unexpectedly lost in the other.

	Nom	Ashé	Paj	Ashá	Kak	Mat
GO	há	ha	ha	ha	ha	a
SPIDER	hito	heto	heto	heto	heeto	eto
FAN (n.)	hobaɾo	heβaa	hewa	heβaɾo		ebaɾo
Loricariidae spp.-2	hétaɾi		hetaɾi	hetaɾi		etaɾi
SKY			henoki	henoki		henokɯi
FAR	háanta	hanto				
discard	hók	ook	hok	hok		ok
HUSBAND	híme	hime	emi	hime		hime
WIFE	hína	hina	ena	hina		hina

Table 81: Retentions of root-initial **h* in Ashéninka and Pajonal

The generalization that **h* was lost intervocalically is supported by reflexes of the fluid classifier **ha* in Pajonal. It is retained word-initially in the words /ha-mento/ (“basin, bowl, plate”) and /ha-meŋkoɾe/ (“raft”), and deleted in words like /ʃiŋkʲ-a/ (“CHICHA, CORN”) and /tʰapʲ-a/ (“shore, bank”).

The following table shows exceptions to the general rule, in which PK **h* was either lost at the beginning of a root or retained intervocalically.

	Nom	Ashé	Paj	Ashá	Kak	Mat
BE CALLED/NAMED	hí	ii	ii	hi		
SPIT (v.)-2	hóbá		eewa	heβ		aba
GOURD VESSEL			paho	pahone	paho	pao

Table 82: Unexpected outcomes of **h* in Ashéninka and Pajonal

4.14.5 Development of **h* in Matsigenka

PK **h* is almost always lost in Matsigenka. I have only five examples in my data in which it was retained. According to personal communication with O’Hagan, the words for “WIFE” and “HUSBAND” may have preserved **h* because they are commonly used kinship terms, and thus more resistant to sound change. I have no satisfactory explanation for the other three retentions.

	Nom	Ashé	Paj	Ashá	Kak	Mat
WIFE	hína	hina	ena	hina		hina
HUSBAND	híme	hime	emi	hime		hime
BAT	pihíɾí	piiɾi	piiɾi	pihiɾi		pihiɾi
BRADYPUS SP.-2					oohi	ohɯi
SKY			henoki	henoki		henokɯi

Table 83: Retentions of **h* in Matsigenka

Another complication is that root-final **h* can resurface in Matsigenka in order to avoid the sequence /aai/. One example occurs in the word /i-tinah-a-ig-an-ak-a/, in which the root meaning “get up” developed from PK **tinah*, as evidenced by the Kakinte reflex /tinah/ (H. V. Pereira and J. V. Pereira 2013, p. 136). Because /h/ only surfaces before specific suffixes that create an /ai/ sequence, and not in the nominalized form from which the roots in my data were extracted, I do not have root-final /h/ in my data. It is unclear if such roots should be analyzed as having underlying /h/, or if forms with /h/ should be seen as prosodically conditioned allomorphs.

4.15 Diversification of Proto-Kampa **ɾ*

In general, PK **ɾ* remains as /ɾ/ in all of the daughter varieties.

4.15.1 Loss of **ɾ* in Pajonal

In Pajonal, PK **ɾ* was lost in *ɑ_ɑ* and *o_o* environments. PK **aɾo* developed into either /awo/ or /aa/, and there are no cognate sets showing the development of PK **oɾa*. Based on these facts, we can conclude that **ɾ* was lost in Pajonal between two non-front vowels, but became the glide /w/ where loss would create an unacceptable vowel sequence. This description is parallel to my description of **g*-loss in Ashéninka and Pajonal in Section 4.9.2. It shifts the question of why we get both ∅ and /w/ in an *ɑ_o* environment into a question about acceptable vowel sequences, which is outside the scope of this article.

	Nom	Ashé	Paj	Ashá	Kak	Mat
Gynerium sagittatum	soboɾo	saβoɾo	hawoo	saβoɾo	saβoɾo	saboɾo
BRADYPUS SP.-1	soɾo	soɾo	hoo	soɾoni		soɾoni
FACE	bóɾo	poɾo	poo	βoɾo		boɾo
HOLE	moɾo	moɾo	moo	moɾo	moɾo	
DRY (vi.)-2			oow	oɾo	oroɣ	oɾog
VOMIT	komaɾaŋá	kamaɾaŋk	kamaaŋk	kamaɾaŋk	kamaɾaŋk	kamaɾaŋk
FLY (V.)	aɾ	aɾ	aaɣ	aɾ	aɾ	aɾ
ORTALIS sp.	maɾáti	maɾatsi	maatsi	maɾati		maɾati
Banisteriopsis caapi	komaɾampi	kamaɾampi	kamaampi	kamaɾampi		kamaɾampi
SNAKE	máɾaŋe	maaŋke	maaŋki	maɾaŋke		maɾaŋke
ARM		naasampi	naahampi			naɾa

Table 84: **ɾ* > ∅ / *ɑ_ɑ, o_o* in Pajonal

	Nom	Ashé	Paj	Ashá	Kak	Mat
DASYPROCTA SP.	ʃaɾo	ʃaɾo	ʃawo			ʃaɾo
ARA MACAO			hawawo	saβaɾo		
granddaughter-2		saβo	hawo	saɾo		pisaɾo
AFRAID	tsoɾóg	t^{h}aaβ	t^{h}aaw		tsaɾoɣ	tsaɾog
OCHROMA PYRAMIDALE	páɾota	paɾoto	paato	paɾoto		paɾoto

Table 85: Development of **aɾo* in Pajonal

The following table shows several instances in which **ɾ* has been retained in an *ɑ_ɑ* or *o_o* environment. These cases may be the result of contact with other varieties of Ashéninka which retain **ɾ*. In the case of /jaaɾato/, it is conceivable that **ɾ* was retained because deletion would create a trimoraic vowel sequence.

	Nom	Ashé	Paj	Ashá	Kak	Mat
Tinamidae spp.-1		saaŋkots^{h}i	ʃaaŋkoɾotsi			
LONTRA LONGICAUDIS	paɾaɾi	paɾaɾi	paɾaɾi	paɾaɾi	βaɾaɾi	paɾaɾi
BEE	eíróto		jaaɾato		eaɾoto	
CERATOPOGONIDAE spp.			joohaɾa			josaɾo

Table 86: Unexpected retentions of **ɾ* in Pajonal

4.15.2 Loss of **r* in Ashéninka

The following table shows cases in which PK **r* was lost or turned into /β/ in Ashéninka. In all but the last two rows, these changes occurred between two non-front vowels, the same environment in which they regularly occur in Pajonal. Thus, the most likely explanation for these cases is that these words in Payne's dictionary were collected from a variety of Ashéninka which has undergone the same process of **r*-loss as the one observed in Pajonal. In the last two rows, PK **r* has apparently been lost in an *e_i* environment. This could be the result of a sporadic change or it could reflect a variety of Ashéninka in which **r*-loss is more widespread than in Pajonal.

	Nom	Ashé	Paj	Ashá	Kak	Mat
granddaughter-1	ʃóɾó	ʃaβo		ʃaβo	ʃao	
granddaughter-2		saβo	hawo	saɾo		pisaɾo
Tinamidae spp.-1		saaŋkots^{h}i	ʃaaŋkoɾotsi			
AFRAID	tsoɾóg	t^{h}aaβ	t^{h}aaw		tsaɾoɣ	tsaɾog
ARM		naasampi	naahampi			naɾa
SNAKE	máɾaŋe	maaŋke	maaŋki	maɾaŋke		maɾaŋke
SPINAL COLUMN	mitíkaɾa	mitsikaa				
BEACH		opaak^{j}a				imparage
Trombiculidae sp.		mampiits^{h}i	mampiɾits^{h}i	mamperitiki	mamperekiti	mamperikiti
CAVE		impeeta				imperitanaki

Table 87: Sporadic intervocalic loss of **r* in Ashéninka

4.16 Diversification of Proto-Kampa **j*

In general, PK **j* remains as /j/ in all of the daughter varieties. I cannot explain the /h/ in Ashéninka /haniɾi/.

	Nom	Ashé	Paj	Ashá	Kak	Mat
PUMA YAGOUAROUNDI	jaíná				jaina	
Tinamidae spp.-2	joŋiɾi			joŋkiɾi		
Brugmansia sp.				hajapa	hajapa	
Swietenia sp.		jopo				jopo
CERATOPOGONIDAE spp.			joohaɾa			josaɾo
BROTH	ijá	jaaki	ijaa			
ALOUATTA SP.	janiɾi	haniɾi				janiɾi

Table 88: Cognate sets showing **j*

4.17 Diversification of Proto-Kampa **N*

In general, PK **N* remains as /N/ in all of the daughter varieties. In Nomatsigenga, **N* merged with /m/ when the following **p* was lost (see Section 4.5.1). Also, **N* developed into the new phoneme /ŋ/ when the following **k* was lost (see Section 4.8.1). The following table shows that PK **N* remained /N/ in *_t*, *_ts*, and *_tʃ* environments in all varieties.

	Nom	Ashé	Paj	Ashá	Kak	Mat
FISH SCALES	bentáki	pentaki	pentaki	βentaki	pentaki	bentaki
ARA ARARAUNA	kasanto	kasanto	kahanto	kasanto		kasanto
PRIODONTES MAXIMUS	kintiɾo	kinteɾo	kinteɾo		kinteɾo	kinteɾo
MOUTH	bagante	paante	paante	βante	βaɣante	bagante
instrumental	-mento	-mento	-mento	-mento	-mento	-mento
unspecified possessor	-ntsi	-nts^{h}i	-nts^{h}i	-ntsi	-ntsi	-ntsi
STRONG	sintsí	ʃints^{h}i	ʃints^{h}i	ʃintsi		ʃintsi
INGA SP.	antsípá	ints^{h}ipa	ints^{h}ipa	intsipa	intsipa	intsipa
HIGH, tall	tsantsáa	sant^{h}a	hant^{h}a	tsantsa	katsantsaheɣi	gatsantsani
LEOPARDUS PARDALIS	matsóntsoɾi	mat^{h}ont^{h}oɾi	mat^{h}ont^{h}oɾi	matsontsoɾi		matsontsoɾi
TREE	aɲtʃáto	iɲtʃ^{h}ato	iɲtʃato	iɲtʃato	iɲtʃato	iɲtʃato
CLOTHING	maɲtʃaki					maɲtʃaki

Table 89: Cognate sets showing **N* in *_t*, *_ts*, and *_tʃ* environments

4.18 Summary of Sound Changes

Here I simply restate the previously described sound changes, but grouped by variety rather than sound. I also include notes on rule ordering if it was discussed above. Time constraints did not permit me to perform a systematic investigation of rule ordering, so these notes are preliminary and by no means intended to be comprehensive.

4.18.1 Nomatsigenga Sound Changes

**h*	> ∅	unclear	
**p*	> ∅	*/ N_*	
**k*	> ∅	*/ N_*	
**g*	> ∅	sporadic	

4.18.2 Ashéninka Sound Changes

**ui*	> *i*		
**b*	> *β*		
**s*	> *ʃ*	*/ _e, _i*	fed by **ui* > *i*
**g*	> *j*	*/ a_e, o_e, e_a, e_i, i_a, i_o, i_e*	
	> *β*	*/ a_o, o_a*	
	> *ɣ*	*/ a_#*	
	> ∅	elsewhere	
**h*	> ∅	*/ V_V*	
**ts*	> *t^{h}*	*/ _a, _o*	
	> *ts^{h}*	*/ _i*	
	> *tʃ^{h}*	*/ _e*	
**tʃ*	> *tʃ^{h}*		
**t*	> *ts*	*/ _i*	after **ts* > *ts^{h}* */ _i* but with overlap, creating some cases of **t* > *ts^{h}* */ _i*
**C*	> *C^{j}*	*/ _ea, _ia, _io*	fed by **h*-loss, after **tʃ* > *tʃ^{h}*, unclear ordering with **g*-loss

4.18.3 Pajonal Sound Changes

**uɨi*	> *i*		
**b*	> *w*		
	> ∅	sporadic	
**s*	> *ʃ*	*/ _e, _i*	fed by **uɨi* > *i*
**g*	> *j*	*/ a_e, o_e, e_a, e_i, i_a, i_o, i_e*	
	> *w*	*/ a_o, o_a, e_o*	
	> *ɣ*	*/ a_#*	
	> ∅	elsewhere	
**h*	> ∅	*/ V_V*	
**ts*	> *t^{h}*	*/ _a, _o*	
	> *ts^{h}*	*/ _i*	
	> *tʃ*	*/ _e*	after **h*-loss
**t*	> *ts*	*/ _i*	after **ts* > *ts^{h}* */ _i*
**C*	> *C^{j}*	*/ _ea, _ia, _io*	fed by **h*-loss, unclear ordering with **g*-loss
**ɾ*	> ∅	*/ a_a, o_o*	
	> *w*	*/ a_o*	

4.18.4 Asháninka Sound Changes

**uɨi*	> *i*		
**b*	> *β*		
	> ∅	sporadic	
**s*	> *ʃ*	*/ _e, _i*	fed by **uɨi* > *i*
**g*	> *j*	*/ _e*	
	> ∅	elsewhere	
**h*	> ∅	unclear	

4.18.5 Kakinte Sound Changes

**uɨi*	> *i*		
**b*	> *β*		
**s*	> *ʃ*	*/ _*i	fed by **uɨi* > *i*
**g*	> *ɣ*		
	> ∅	unclear	
**t*	> *t^{j}*	sporadic	
**n*	> *ɲ*	sporadic	

4.18.6 Matsigenka Sound Changes

**s*	> *ʃ*	*/ _i*	
**t*	> *t^{j}*	*/ _iV*	before **h*-loss
**h*	> ∅		
**n*	> *ɲ*	sporadic	

5 Conclusion

In this section I summarize my research, reflect on the computational tools that I used, discuss evidence for internal subgrouping of the Kampa family, and suggest areas for further study.

5.1 Summary

In this article, I presented the first comprehensive reconstruction of Proto-Kampa consonants using the comparative method. I used a suite of computational tools developed by Johann-Mattis List to improve the efficiency of my work. Particularly notable is the fact that this is possibly the first reconstruction to use LingRex, a tool for automatic generation of correspondence sets. Using these correspondence sets, I reconstructed the Proto-Kampa consonant inventory and argued for its correctness in Section 3. In Section 4, I laid out the various sound changes that led to the diversification of Proto-Kampa consonants in the daughter varieties, noting exceptions and unexplained phenomena that require further study.

5.2 Evaluation of Computational Tools

In my research, I used two Python packages for computational historical linguistics: LingPy provided algorithms for automatic cognate detection and alignment of cognate sets, and LingRex provided an algorithm for automatic generation of correspondence sets. I also used Edictor, a browser-based app which allowed me to view and edit the outputs of LingPy and LingRex, among other features. All of these tools were created by Johann-Mattis List.

On the whole, these tools saved me an inestimable amount of time in the process of generating correspondence sets from wordlists. There is a slight overhead in preparing the data into a format understood by LingPy as well as in learning to use the tools (provided that one has some basic knowledge of programming). This is more than outweighed by the fact that LingPy and LingRex work on the order of minutes, with a level of accuracy that makes the revision process not too tedious.

Given the limitations on LingPy and LingRex, caused by the fact that we cannot yet encode certain types of linguistic knowledge into an algorithm, they were about as accurate as I could have hoped for (see List's book for a quantitative evaluation of LingPy (List 2014, p. 185)). However, we are still quite far from the goal of automating the entire reconstruction process. I expect great advances to be made in this regard as the field of computational historical linguistics continues to progress.

Several improvements could be made to Edictor in order to streamline the user experience. At times, a panel in Edictor or the whole program would crash due to some error in the underlying spreadsheet. It would be useful for Edictor to provide error detection to inform the user where corrections need to be made. Also, as LingRex is a very new tool, Edictor does not have a feature for editing correspondence sets like the panels for editing cognate sets and alignments. It would be useful to be able to change the correspondence set of an entire site at the same time (i.e. a given column of Figure 6, rather than one cell at a time).

5.3 Evidence for Subgrouping

As I discussed in Section 1.4.2, previous work on the internal subgrouping of the Kampa family has identified two candidate structures, shown in Figures 2 and 3. My reconstructed sound changes are consistent with both of these structures, but do not provide any evidence favoring one over the other.

First, it is obvious that Ashéninka and Pajonal form a subgroup among the varieties in my sample. Then, there is some evidence that Asháninka, Ashéninka, and Pajonal form a subgroup. Possible shared innovations among this subgroup are **s* > *ʃ* / _*e* and lenition or loss of **g*. However, both of these sound changes are cross-linguistically quite common, so they could have occurred independently. Also, positing changes to **g* as a shared innovation is somewhat problematic because **g* develops differently in Asháninka compared to Ashéninka and Pajonal (see Section 4.9), so it is unclear what change should be considered the shared innovation in the common ancestor. Furthermore, both **s* > *ʃ* / _*e* and loss of **g* are attested in Kakinte, so they are not great evidence for Asháninka, Ashéninka, and Pajonal forming a subgroup to the exclusion of Kakinte.

There is also some evidence that Ashéninka, Pajonal, Asháninka, and Kakinte form a subgroup. Possible shared innovations are **b* > *β* and **ɯi* > *i*. Again, these changes are not particularly

uncommon, so they could have occurred independently in separate branches of the family.

Evidence for a shared innovation in all varieties except Nomatsigenga would support the structure in Figure 2. Michael previously mentioned **s* > *ʃ / _i* as a possible candidate, but I showed in Section 4.12.1 that it is unlikely to be a shared innovation in all non-Nomatsigenga varieties. Evidence for a shared innovation in Nomatsigenga, Matsigenka, and Nanti would support the structure in Figure 3, but there is no clear candidate at least in the realm of sound changes affecting consonants. More work in the reconstruction of vowels or morphology may shed light on this issue.

5.4 Areas for Further Study

Here I suggest avenues for further research regarding the history of the Kampa family. First, as discussed throughout Section 4, I do not have a satisfactory explanation for various irregularities relating to consonant sound changes. The following problems are especially salient:

1. What factors explain the loss or retention of **g* in Kakinte?
2. What is the best explanation for the development of **g* in Ashéninka and Pajonal? A better understanding of the development of vowels is likely necessary to answer this question.
3. What is proper synchronic analysis of /h/ and processes of /h/-loss in the modern Kampa varieties? How can these processes help us understand the historical process of **h*-loss?
4. What factors explain the loss or retention of **r* in Ashéninka? Better documentation of Ashéninka varieties is necessary to answer this question.
5. What is the history of the palatalized consonants in Kakinte and Matsigenka? Should the development of palatalized consonants in Kakinte and its lexical divergence be understood as an instance of esoterogeny?
6. What factors explain the retention of **Np* and **Nk* sequences in Nomatsigenga? Are all of these retentions the result of contact with neighboring varieties?
7. Is it true that verb roots ending in **Ceh* and **Cih* palatalized in Ashéninka and Pajonal, but those ending in **Ce* and **Ci* did not?

Beyond the reconstruction of consonants, the obvious next step is the reconstruction of vowels and associated sound changes using the comparative method. Ideally, this research would draw on not only the six Kampa varieties studied in this article, but also Nanti and other varieties of Ashéninka. A full reconstruction of Proto-Kampa phonology would discuss issues of rule ordering and interactions between vowel changes and consonant changes, and may provide evidence elucidating the internal structure of the Kampa family. It would also bring us much closer to the goal of reconstructing the Proto-Kampa lexicon.

Of course, a fuller understanding of Proto-Kampa would also require morphological and syntactic reconstruction. I believe that more detailed and comprehensive synchronic analyses of Kampa varieties is needed before we can progress in these aspects of reconstruction.

References

Aikhenvald, Alexandra Y. (1999). “The Arawak language family”. In: *The Amazonian languages*. Ed. by Robert M. W. Dixon and Alexandra Y. Aikhenvald. Cambridge, United Kingdom: Cambridge University Press, pp. 65–106.

Clem, Emily and Lev Michael (2015). “Exploring typological diversity in South America and its genealogical and areal basis”. Arealphonologie workshop. Bremen, Germany.

Coward, David and Charles E. Grimes (2003). *Making dictionaries: a guide to lexicography and the Multi-Dictionary Formatter*. Dallas, Texas: SIL International DigitalResources. URL: http://downloads.sil.org/legacy/shoebox/MDF_2000.pdf.

Dimmendaal, Gerrit Jan (2015). *The Leopard's Spots: Essays on Language, Cognition and Culture*. Brill. DOI: 10.1163/9789004224148_005.

Dohn, Michael (2017). "Directionals and Aspect in Matsigenka".

Heitzman, Allene R. (1973). "Reflexes of Some Proto-Campa Consonants in Modern Campan Languages and Dialects". MA thesis. Waco, Texas: Baylor University.

Heitzman, Allene R. and Verna Doerksen (2017). "El alfabeto asheninca pajonalino".

— (2018). "Diccionario asheninka pajonalino (manuscript)".

Kindberg, Lee D. (1980). *Diccionario Asháninca: Edición Provisional*. Yarinacocha, Peru: Instituto Lingüístico de Verano.

Lawrence, Aimee Lynn (2012). "Reconstruction of Proto-Kampa Verbal Morphology". In: *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*. DOI: 10.3765/bls.v38i0.3334. URL: https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/3334.

— (2013). "Inflectional Verbal Morphology in Nomatsigenga". MA thesis. Austin, Texas: University of Texas at Austin.

List, Johann-Mattis (2014). *Sequence Comparison in Historical Linguistics*. Dissertations in Language and Cognition 1. Düsseldorf, Germany: düsseldorf university press. ISBN: 978-3-940671-72-8. URL: http://dup.oa.hhu.de/244/1/DLS%20(J.-M.%20List)%2C%20Vol.%201_Open%20Access.pdf.

— (2017a). "Historical language comparison with LingPy and EDICTOR". In: DOI: 10.5281/zenodo.1042204. URL: https://github.com/digling/edictor-tutorial.

— (May 10, 2017b). *Sequence Comparison with LingPy*. URL: http://lingulist.de/documents/tutorials/list-2017-sequence-comparison-lingpy-tutorial.pdf.

— (2018a). *EDICTOR*. URL: http://edictor.digling.org.

— (2018b). *LingRex: Linguistic reconstruction with LingPy*. Jena. URL: https://github.com/lingpy/lingrex.

— (2019). "Automatic Inference of Sound Correspondence Patterns Across Multiple Languages". In: *Computational Linguistics*. DOI: 10.1162/coli_a_00344. URL: https://www.mitpressjournals.org/doi/full/10.1162/coli_a_00344.

List, Johann-Mattis, Simon Greenhill, and Robert Forkel (2017). *LingPy. A Python library for quantitative tasks in historical linguistics*. Jena. DOI: 10.5281/zenodo.1065403. URL: http://lingpy.org.

Michael, Lev (2008). "Nanti evidential practice: Language, knowledge, and social action in an Amazonian society". PhD thesis. Austin, Texas: University of Texas at Austin. URL: https://repositories.lib.utexas.edu/handle/2152/3884.

— (2011). "La reconstrucción y clasificación interna de la rama Kampa de la familia Arawak". CILLA V.

Michael, Lev et al. (2010). "Phonological Reconstruction of the Kampan branch of Arawak". Workshop on American Indian Languages.

Mihas, Elena (2017). "The Kampa subgroup of the Arawak language family". In: *The Cambridge Handbook of Linguistic Typology*. Ed. by Robert M. W. Dixon and Alexandra Y. Aikhenvald. Cambridge, United Kingdom: Cambridge University Press, pp. 782–814.

O'Hagan, Zachary (2016). "Comparative Kakinte Lexical Data".

Payne, David Lawrence (1980). *Diccionario Asheninca-Castellano: Edición Preliminar*. Yarinacocha, Peru: Instituto Lingüístico de Verano. URL: https://www.sil.org/resources/archives/30268.

— (1981). *The Phonology and Morphology of Axininca Campa*. Ed. by Virgil L. Poulter. Summer Institute of Linguistics and University of Texas at Arlington. URL: https://www.sil.org/resources/archives/8521.

Pedrós, Toni (2018). "Asheninka y asháninka: ¿de cuántas lenguas hablamos?" In: *Cadernos de Etnolingüística* 6.1, pp. 1–30. URL: http://etnolinguistica.wdfiles.com/local--files/cadernos%3Avol6n1/cadernos_de_etnolinguistica_vol6_no1.pdf.

Pereira, Haroldo Vargas and José Vargas Pereira (2013). *Matsigenka Texts*. Ed. by Lev Michael, Christine Beier, and Zachary O'Hagan. URL: http://linguistics.berkeley.edu/~zjohagan/pdflinks/mcb_text_collection_30jun2013_v1.

Shaver, Harold (1996). *Diccionario Nomatsiguenga-Castellano Castellano-Nomatsiguenga*. Ed. by Mary Ruth Wise. Yarinacocha, Peru: Ministerio de Educación and Instituto Lingüístico de Verano. URL: https://www.sil.org/resources/archives/29591.

Snell, Betty (2011). *Diccionario Matsigenka-Castellano*. Ed. by Mary Ruth Wise. Lima, Peru: Instituto Lingüístico de Verano. URL: https://www.sil.org/resources/archives/43040.

Swift, Kenneth E. (1988). *Morfología del Caquinte (Arawak Preandino)*. Ed. by Mary Ruth Wise. Yarinacocha, Peru: Ministerio de Educación and Instituto Lingüístico de Verano. URL: https://www.sil.org/resources/archives/29899.