A New Approach to the Reconstitution of the Pronunciation of Timote-Cuica (Venezuelan Andes)

Matthias Urban
Center for Advanced Studies 'Words, Bones, Genes, Tools'
University of Tübingen

Abstract

This article concerns the extinct and poorly described native speech of the Venezuelan Andes conventionally known as Timote-Cuica, in particular its phonetics and phonology. While the available pre-phonemic and unsystematically transcribed corpus of data (consisting of about 900 words and 300 phrases and sentences) has already been analyzed using the method of reconstitution of imperfect data, remarks by the transcribers on the sound of Timote-Cuica have not been taken into account so far. Here, it is shown that these provide valuable clues to the reconstitution of Timote-Cuica pronunciation. In particular, such observations in conjunction with a close analysis of the available data reinforce the idea of the presence of a high central vowel, but also suggest hitherto unrecognized properties, notably the presence of prenasalized stops that likely contrasted with their plain counterparts.

1 Introduction

As the extinction of the world's languages is continuing at an unprecedented pace, in the case of many languages it is already too late for documentary efforts. As elsewhere, indigenous peoples of the Americas and their languages have come under extraordinary pressure ever since the European conquest. Driven by devastating population losses, it must in some cases have been within only very few generations of speakers that the language shift to Spanish was complete in. This scenario, for instance, probably applied to Quingnam, a once important language of coastal Peru. Other peoples under different circumstances managed to maintain their linguistic identity for a significantly longer period but unfortunately yet not to the present day.

Such is the case with the group of doculects (see Cysouw and Good 2013 for the term) this contribution is concerned with. Conventionally known under the label Timote-Cuica, they are usually considered to form a small linguistic family with hitherto unclear further external connections. [1] Natively spoken in the Venezuelan Andes (principally Mérida and the western parts of Trujillo), speakers are known to have lived into the 20th century and rumors suggest at least the possibility that the language is conserved until the present day in remote villages (Adelaar with Muysken 2004:125). Until such reports can be confirmed, Timote-Cuica must be considered extinct. Given the time of the presumed extinction, it was, as in the case of many other languages, too early for data to be gathered in the context of modern linguistic fieldwork. What data we have comes down to us in unsystematic transcriptions, collected by local intellectuals without training in phonetics. Such data seeks, surely not always successfully, to represent raw phonetic surface data using an alphabet that is likely to have been inadequate for the representation of the phonetic structure of the target language. We will explore the properties of this data in more detail in the following section.

---

[1] I stick with the most established label Timote-Cuica for the purpose of this article but would like to point out that Timote-Cuicas may actually be more appropriate. This is because Cuicas may be composed of an amply attested article <kiu> ~ <ki> and a root for 'Indian person, man' which is among other ways transcribed as <kas> by Fonseca (2005) and the final consonant of which may have later been erroneously identified as part of the Spanish plural morpheme and stripped off to yield Timote-Cuica.

While many linguists will feel a strong urge to keep their hands off these kinds of materials, and indeed unless for specific philological purposes (cf. Goddard 1973) no one would want to use these as the principal linguistic source, those who wish to learn anything about the languages have no choice other than to go through such data. And there are good reasons for not bypassing them: along with ethnohistorical evidence, they are important building blocks in the enterprise of reconstructing the linguistic landscape of South America as it existed at the point of European contact. That landscape, in turn, not the linguistic distribution in the present day, should form the basis for theorizing the continent's original linguistic ecology. Especially if the documented languages appear to be isolates or pertain to small families, as is the case here, the understanding of early sources should be as exhaustive and detailed as possible because only if the data are understood avenues for meaningful external comparison have the potential to become available. And a close analysis can also shed light on more local linguistic ecologies and patterns of language contact, as the presence of numerals with an Arawakan source in the Timote-Cuica materials demonstrates (Urban 2015).

2 The Available Data

Not counting repetitions of the same items, the corpus of Timote-Cuica data consists of approximately 900 form-meaning pairs ("words") and 300 phrases and sentences, most just attested for a single doculect. Unfortunately, the history of the extant data is extremely complex and only partially reconstructible. The peculiarities of this history also make it difficult to decide if Timote-Cuica is best considered a small language family or a dialectally diversified single language. The following are, in broad strokes, key obvervations which are indispensable for proper interpretation of the data and for understanding why the status along the dichotomy of dialect vs. distinct languages is unresolved.

The history of principled efforts at documentation of the indigenous languages of the Venezuelan Andes appears to begin with the work of Rafael María Urrecheaga. Born in 1826 in La Mesa de Esnujaque in the state of Trujillo close to the present-day border to Mérida (de Santiago 1956:299), Urrecheaga recuperated linguistic material from the indigenous people of his home region between 1875 and 1878 (Jahn 1927:325). Urrecheaga then sent copies of the fruits of his work, 16 sheets of notes in two notebooks (Pérez Hernández 1988:160) to Arístides Rojas, of Caracas (Chamberlain 1910:199). Urrecheaga had planned to publish his data (under the title "Noticias sobre la lengua de los Timotes" according to Briceño-Iragorry 1929:156) but died in 1894 before being able to do so. Alfredo Jahn, himself interested in the ethnography and linguistics of western Venezuela, was not able to find the original manuscript in Rojas's library, but finally obtained a copy from Amílcar Fonseca, who, as Jahn reports, was in possession of the original manuscript at the time of writing. Even though lost, Urrecheagas work forms the basis of available work on the variety conventionally known as Cuica.

According to Jahn (1927:326) himself, Urrecheaga's manuscript served as basis for the vocabulary contained in his own work. Jahn reordered Urrecheaga's data and extracted words from phrases to create a vocabulary. As late as 1910 Jahn encountered in La Mesa de Esnujaque two old indians who knew the language, and with their help he double-checked and amended Urrecheaga's data himself (Jahn 1927:327). Unfortunately, in spite of the fact that Jahn (1927:326) explicitly states that the orthography he employs is the same as Urrecheaga's, the possibility that Jahn altered either of the datasets to make them consistent, thereby levelling important clues as to original pronunciation and grammatical structure, is very real.

Amilcar Fonseca, from whom Jahn had finally obtained the Urrecheaga manuscript, published the data and his interpretations in the form of short articles in various Venezuelan periodicals in the 1910s and 1920. Rivet (1927) procured and evaluated many of these articles. Rivet (1927:142-143) also mentions that they appeared to be fragments of a series that together would form a single longer study of three chapters, each featuring an appendix and a vocabulary. Indeed, Fonseca's monograph Orígines Trujillanos, republished in two volumes in 2005, features a chapter entitled 'Dialectología' on pp. 261-338 of volume 1, which seems to be the coherent study Rivet suspected. This compilation is used as the source for the present study.

Nowhere in this chapter does Fonseca indicate that his data actually come, at least in part, from Urrecheaga, but this seems beyond reasonable doubt. Both Jahn's account of the history of the data as well as the

similarities in the recorded phrases in both sources suggest that much. Briceño-Iragorry (1929:157) states that Fonseca did amend Urrechuega's materials personally in Santana (Trujillo district), but the precise circumstances and extent of this work are not known.

Given that both Jahn and Fonseca rely on the Urrechuega manuscript (which is also the reason why Adelaar with Muysken 2004:124 note strong similarities between their data), they cannot count as independent witnesses of Cuica. However, since Jahn, and also Fonseca if Briceño-Iragorry is correct, reworked the Urrecheaga data independently, any difference between their work (e.g. items appearing in one but not the other source) is of special interest, since these are potentially independent of one another.

The so-called Timote data from the state of Mérida have a similarly complicated history. Data collected by José Ignacio Lares at various localities were partially published by Lares (1907; 1918) himself, but in a more complete manner in Ernst (1885; 1886) and in Calcaño (1886). In Ernst (1885), each dataset bears the heading "Dialekt" (dialect) followed by the name of the locality. There is a wordlist of 11 items from Mirripú, about 50 words and phrases from Mocochies, about 30 words and phrases from Migurí, three forms labeled as pertaining to "Mocochis von Torondoy", two forms from Tiguñío and Escagüey each, and eight forms from Timotes. Ernst (1886) contains around 50 more words and phrases from Mirripú. Not all localities can be identified unambiguously. The data published in Calcaño (1886) differ somewhat in spelling of the localities as well as the linguistic forms themselves. Calcaño (1886) also features a vocabulary of the "dialecto de los indios de los Andes" which has no equivalent in Ernst (1885). As Rivet (1927:141) already noted, this appears to be an extract from the Migurí vocabulary, to which it is, the differences in spelling and the differences attributable to typesetting errors aside, identical. In the vocabularies in Lares (1907:23-30), there are several likely errors to which Rivet (1927:141-142) has already drawn attention. In the following, I use the following standardized designations for the various wordlists collected by Lares, usually taken as samples of the Timote language of Mérida:

Standardized	Lares in Ernst (1885)	Lares in Ernst (1886)	Lares in Calcaño (1886)	Lares (1907)	Lares (1918)
Mucuchíes	Mocochies	-	Mocochís	Mucuchís	Mucuchíes
Mirripú$_1$	Mirripú	-	Mirripú	Mirripú	Mirripú
Mirripú$_2$	-	Mirripú	Mirripú (?)	-	-
Migurí	Migurí	-	Migurí	Migurí	-
Torondoy Mucuchíes	Mocochis von Torondoy	-	Mocochis del Torondoy	Mucuchís del Torondoy	Torondoy
Tiguñió	Tiguñió	-	Tiguñió	Tiguñió	-
Escagüey	Escagüey	-	Escaguey	Escaguey	-
Timotes	Timotes	-	Timotes	Timotes	-

Table 1: Standardized designations for Timote doculects in José Ignacio Lares's work

The data of another Timote scholar, Tulio Febres Cordero, were first published in Ernst (1885) and just shortly after in Calcaño (1886). The data in Ernst (1885), a list of 37 items, are labeled as pertaining to the dialect of El Morro. The data in Calcaño (1886) consist of a list of around 60 items labeled as "Dialecto Maripuyes de 'El Morro'", which is only partially overlapping with the list in Ernst (1885) and, unlike the earlier list, feature the numerals up to 23. I use the standardized designations "El Morro Maripuyes" and "El Cenicero Mucuchíes" as specified in table 2.

	Febres Cordero in Ernst (1885)	Febres Cordero in Calcaño (1886)	Febres Cordero in Calcaño (1886)
El Morro Maripuyes	El Morro	Dialecto Maripuyes de "El Morro."	-
El Cenicero Mucuchíes	-	-	Dialecto Mucuchies de "El Cenicero"

Table 2: Standardized designations for Timote doculects in Tulio Febres Cordero's work

Another quite extensive list of vocabulary, accompanied by valuable notes on the circumstances of data collection and on the pronunciation of the language(s) that will be discussed extensively in this article, appears in Febres Cordero (1921). Febres Cordero (1921:iii) states that his vocabulary pertains to the Mucuchíes and Mucubaches varieties. He says that from 1884 onward he personally collected vocabulary in Mucuchíes and the villages of southern Mérida. He also says that his vocabulary includes those published by Lares as well as items he procured from yet other individuals, namely Salvador Dugarte (whose data are from El Morro) and Balbino Balza (data from Mucuchíes) (Febres Cordero 1921:113). Unfortunately, this means another complication which makes it next to impossible to identify a transcriber for a given form in Febres Cordero (1921). One cannot simply subtract the forms from Lares's and Febres Cordero's data which are published elsewhere and conclude that everything else was collected by Dugarte or Balza, since Febres Cordero may have proceeded with his work on the languages since his first publications from the 1880s and added new items collected by himself. I therefore treat Febres Cordero (1921) as a separate source.

In sum, the available Timote-Cuica data consist of (i) the Urrecheaga-based Cuica data of Jahn (1927) and Fonseca (2005) and (ii) the various wordlists for Timote as just described. The complex history of the data is all the more unfortunate as the degree of independence between the sources is correlated with the commonly made division between Timote (with data collected mainly by Febres Cordero and Lares) on the one hand and Cuica (with data from Jahn and Fonseca on the basis of Urrecheaga) on the other. In fact, however, neither the data published by Fonseca and by Jahn on the one hand nor those by Febres Cordero and Lares on the other can be taken uncritically as independent witnesses. In the absence of data in which one and the same transcriber provides data for both the lects of Mérida ("Timote") and Trujillo ("Cuica"), we cannot properly distinguish between variation which is due to different transcription habits that became established among the two groups of scholars and actual linguistic variation that might help to decide the question as to whether what is known as Timote-Cuica was indeed a small language family rather than a single language with diatopic variation. The documentary history might cement an ethnolinguistic distinction between Timote and Cuica which may or may not have been artificial in the first place on linguistic grounds. Accordingly, "Timote" and "Cuica" are more appropriately considered labels for data coming from these research traditions rather than necessarily labels for languages or dialects which are clearly distinguishable on linguistic grounds.

However, this does not mean that comparison between transcriptions of individual transcribers cannot be insightful: each independent transcription of words is important for reconstituting the described language, because with each independent transcription (preferably by transcribers with different native languages, which is not the case here either) that can be added to the pool of available data, the original sounds heard by the transcribers and the shape of the described language's words can be recovered more securely (Broadbent 1957). Of special interest are also words that appear only in one of the respective sources, because these words can, with relative security, be ascribed to the author who is also actually credited as the origin of the data. Therefore, I will keep the data from the four scholars –Jahn and Fonseca for Cuica and Lares and Febres Cordero for Timote–separate from one another, though throughout it needs to be borne in mind that mutual influence probably took part to some extent.

3 The Methodology of Reconstitution and the Value of Naïve Observers' Auditory Impressions

How is such material best approached? Apparently independently, two scholars have suggested highly similar principled methods to make the most out of asystematically transcribed material such as that available for Timote-Cuica.

Broadbent's (1957) method, called "reconstitution", grew out of an attempt to learn more about the sound structure of Rumsen, an extinct Californian language, than the nine vocabulary lists from different authors from the 18th to the 20th century which make up the entire corpus of available material reveal at first hand. Broadbent explicitly describes the steps involved in the method; I merely summarize the most important points here. In order to "try to find out what the forms actually sounded like" (1957:277), one

arranges the available material according to English glosses so that different transcriptions of the same form appear together. These are then compared form by form, starting with such items for which more than one transcription is available. In the case that there is agreement between different transcribers on the sound recorded, this is assumed to be "reasonably secure." Where differences exist, a likely pronunciation is reconstituted on the basis of the emerging alternations. Here, Broadbent's recommendation to use material as close as possible to the original recording situation becomes important. Also, the native language of the transcribers and their phonetic and phonological structure must be borne in mind in this situation. Broadbent emphasizes that "[t]he process is not a simple mechanical one; each case must be carefully considered and decided on its own merits" (1957:279). Broadbent's instruction unfortunately presents rather few actual examples from Rumsen, but her example of reconstituting [p] for a sound written by Spanish and French authors as <p> and by English authors as <b> gives an idea of the kind of reasoning that must be involved here. Another example of reconstituting a sound involves a single transcriber vary in their spelling of the same word in an apparently random manner, a situation which suggests to Broadbent that the transcriber was unable to pin down the quality of the sound and that what was actually heard was most likely somewhere in between the canonical values of the alternating letters. Once one has arrived at a likely phonetic representation of the forms, one can apply principles of phonological analysis to phonemicize the raw data in order to learn about the sound patterns and the phoneme system of the analyzed language. Broadbent suggests the term sone and soneme, analogous to phone and phoneme, for the reconstituted units.

Constenla Umaña (2000) offers another elaboration of a principled approach to linguistic data available only in imperfect transcriptions. The name Constenla Umaña gave to his technique, "restitución", and the close similarities of his approach with Broadbent's is very striking given that it was apparently developed independently. As far as the establishment of the phonetic value of the letters employed by transcribers, Constenla Umaña (2000:165-167) provides more ample discussion of ancillary factors that may give clues about the sound values of the letters. Specifically, he mentions the nature of the employed alphabet for transcription, the era from which the data stem (since different letters may have represented different sounds at different times), local traditions in transcribing certain sounds with certain letters, personal characteristics of the (hand-)writing of the transcribers, observations made on behalf of the transcriber or others on the qualities of the sounds in the language, characteristic patterns in the adaptation of loanwords, comparison of transcriptions in the case that a letter in the employed alphabet can have more than one phonetic value, and finally, knowledge of the areal-typological characteristics of the sound systems of other languages in the same region as well as the sound systems of genealogically related languages in case such exist.

In fact, the Timote-Cuica data have already been reconstituted by Arrieta (1993; 1998), a student of Constenla Umaña's. Many interesting features of Timote-Cuica phonology and morphosyntax have been unveiled by her analysis. In spite of the fact Constenla Umaña explicitly urges to take into account observations on the phonetic quality of sounds or the gestures involved in their articulation, however, Arrieta does not seem to have taken into account valuable statements by Febres Cordero (1921). Also her predecessor Rivet (1927), although he quotes from Febres Cordero (1921), bypassed these important comments, since he seems to have been less interested in phonetics and phonology but more in morphosyntax. In fact, Rivet generously replaced all instances where <j> would have been pronounced [x] or [h] by <x>. This can be extraordinarily harmful in the case of languages whose actual phonetics and phonology are not known, since there is a very real chance that the symbols sought to represent something other than fixed qualities associated with the letters in a conventional orthographic system. In this article, then, I set out to evaluate the data anew, paying closer attention to transcription habits of individual sources, but especially to observations by those who actually heard the language when it was still spoken.

To be able to link statements to actual transcriptions later, I have created a spreadsheet containing all available Timote and Cuica data from the sources mentioned in section 2. Items with the same meaning and similar form (i.e. likely cognates) were placed into the same row to facilitate comparison. The wordlists from the different localities and authors are put into distinct columns so that comparisons can easily be carried out also across authors.

4 Vowels

4.1 Vowel Qualities

I begin the evaluation of observer's statements on the pronunciation of Timote-Cuica with vowels and will move to consonants in the following section. About his perception of the vocalic sounds, Febres Cordero (1921:92) remarks:[2]

> Existen dos vocales que no pueden expresarse en castellano, una que equivale a un sonido compuesto de *o* y *u*, como en la voz *mucu* o *moco* ; y otra que equivale a uno intermedio entre la *i* y la *u*, el cual se advierte en *capsín* o *capsún*, que de ambos modos puede escribirse, según la varia apreciacíon de quien oye a los indios.

> There are two vowels which cannot be pronounced in Spanish, one which corresponds to a sound composed of *o* and *u*, like in the word *mucu* or *moco*, and another which corresponds to one in between *i* and *u*, which one notes in *capsín* or *capsún*, which can be written in both ways, corresponding to the varying assessment of the one who hears the Indians.

Febres Cordero goes on to note that the former vowel, the one "composed of *o* and *u*", is found in many languages of the Americas, among them Huron and Iroquoian as well as Nahuatl. Salas (1908:77) also mentions difficulties on behalf of the Spanish in pronouncing the placenames of the Venezuelan Andes with the characteristic mucu/moco element. According to him, they feature a sound which is "guttural and nasal at the same time, oscillating between mu, mo, and mgo" ("un sonido gutural y nasal al mismo tiempo, que oscila entre mu, mo y mgo")

From an articulatory point of view, the best candidate for a sound "composed of *o* and *u*" is [ʊ]. This sound shares the rounded quality of [o] and [u], and indeed, tongue height is intermediate between [o] and [u]. A sound "between *i* and *u*", on the other hand, makes one think of the central high vowels [ɨ] and [ʉ]. Also Jahn (1927) makes reference to two vocalic sounds which do not occur in Spanish. According to him, they are pronounced like German <ö> and <ü> (i.e. [ø] and [y]), and he uses the respective letters in his data. However, the frequency of these letters is quite low.

It is perhaps significant that in one case Febres Cordero speaks of a vowel "composed of *o* and *u*" and in the other of one "between *i* and *u*." This choice of wording may point towards a diphthong or a vowel with an onglide or offglide phase. In the Urrecheaga-based Cuica data of Jahn and Fonseca, <eu> (<eú>, <éu>) is a digraph that occurs relatively frequently without notable distributional restrictions, occurring in a wide range of variants. It therefore seems promising to investigate how Timote vowels in the positions corresponding to those where the conspicuous Cuica transcriptions are transcribed.

Table 3 presents those items in Jahn where the sequence <eu> (with variants <eú> or <éu>) or <ö> or <ü> occurs which have equivalents in the Méridan varieties of Timote as recorded by Febres Cordero and Lares.

This yields the correspondence sets in table 4 for <eu> and its variants and those in table 5 for <ö>. These reveal patterns in the data that are as regular as one can reasonably expect in the unsystematically transcribed data at hand: With very high consistency, Jahn's <eu> and variants for Cuica corresponds to <u> in the Timote data. Sometimes there is an alternant with <i> at the end which is in some instances also suggested in the likewise Urrecheaga-based data of Fonseca, e.g. when he transcribes <tigüis>, in others they may reflect a genuine property specific to Timote speech. It is also relevant that <kapsín> is precisely the item of which Febres Cordero says that it can also be written <kapsún>. And of course, also in the Timote

---

data we occasionally find sequences like <ui> or <iu> where in Cuica Jahn and Fonseca chose <eu> (or a variant thereof). All this suggest a connection –whether it is merely orthographic or represents a regular sound correspondence between Timote and Cuica– of <eu> and variants in the Urrecheaga-based data and the sound "between i and o" described by Febres Cordero (1921). Given the alternations it is involved in the correspondence sets, the digraph <eu> is a quite plausible candidate for the orthographic representation of the sound in question which indeed could have been the high central vowel [ɨ], possibly with an onglide phase particularly in Cuica.

    What the data also show is that Jahn's <ö> in Cuica has quite different corresponding sounds when compared with <eu> and variants. In the case of <ö> they appear to cover the entire high and mid vowel space. Even though this makes it very difficult to specify the nature of the original sound that was heard, it should make the interpreter wary of assigning the correspondence set lightheartedly to the ones of <eu>. Whatever phonetic reality is hiding behind <ö> and its counterparts, it likely was a different one from that of <eu>.

4.2 Nasal Vowels

Arrieta (1998:89) notes that <m> and <n> in syllable-final position often alternate with zero. In her orthography, she mentions the following examples:

(1)   a.   *shimpué ~ shipué*   'water'
       b.   *funt ~ xut*   'scorpion'
       c.   *keún ~ kuu*   'sleep'
       d.   *estuce ~ ismtuch*   'flame'
       e.   *xok ~ köngk*   'nose'
       f.   *tump ~ tuup*   'stone'
       g.   *sharin ~ chivatsari*   'close'

She interprets this alternation as indicative of nasalization of the preceding vowel, an interpretation that entails that /m/ and /n/ were contrastive only in the initial position. The existence of nasal vowels is in fact significantly supported by a casual comment by Lares (1907:26). Lares remarks that in the phrase <machini-pe in>? 'who is coming?', the sequence <in> is "entirely nasal" ("*enteramente nasal*"). However, as we will see in the following section on consonants, the interpretation of <m> and <n> as always indicating nasal airflow in the preceding vocalic segment is problematic given that the same letters also are involved in the representation of prenasalization in a significant number of instances.

	Jahn	Fonseca	Febres Cordero (1921)	Mirripú₂	Migurí	Tiguñió	Mucuchíes	El Morro Maripuyes	El Cenicero Mucuchíes
'six'	\<katséun>	\<catseunt>, \<katseun>	\<cacsún>	\<kasum>	\<cassúm>, \<kassum>			\<kapsín>	
'feet'	\<tit-kuaéu>, \<kiéu>	\<kuahem>, \<ti kuaeu>, \<tit kuaes>,\<tit kuaeu>	\<cujú>, \<quijuy>				\<cujú>, \<kujú>		
'water'	\<shömbuch> ~ \<shömbeuch>	\<shombeuch>, \<sombeuch>, \<shombuch>	\<chimbú>, \<chimpú>, \<chumpú>	\<chimpu>	\<chimbú>		\<chimpué>, \<chimpiú>, \<chum piú>		\<chumpuk>
'stone',	\<kiu-teunch>		\<tuch>, \<tump>	\<tituup>					\<tuch>
'louse'	\<kiu-tiéüi>	\<ti tieui>	\<guy>	\<tigui>, \<tiguy>					
'potato'	\<kiu-tiguéus>	\<tigueus>	\<buis>, \<guis>, \<esticuís>, \<ti-guís>, \<tinguís>, \<tigurús>, \<tigús>	\<tigurus>	\<tiguss>		\<tiguis>	\<tingüi>, \<tingüis>	
'fire'	\<ki-shnöpa>, \<shnöp>	\<shnopa>, \<sno-pa>, \<tit shnopá>	\<chirup>, \<sirup>	\<chirup>	\<chirúp>, \<chirup>			\<karí sirup>	\<churup>
'wind'	\<ki-huéreuch>, \<huer kúch>	\<heuur kuch>, \<heurscuch>, \<keurcuchi>, \<keurkuch>, \<heu-rkuch>	\<esfuche>						
'cow, milk'	\<kiu-shichió>, \<tshu-chiök>	\<tchuchio>	\<schó>						
'firewood, strength'	\<ki-tishép>, \<shüshöp>, \<kiu-shushöp>, \<tishöp>, \<tisép>, \<shusép>, \<tishép>	\<tishep>, \<tíshep>, \<ti shuopa>		\<tiscep>, \<tisseph>	\<ti-semp>			\<timpuce>, \<timpuek>	
'hunger'	\<nakshöm>, \<naksöm>,					\<som>			

Table 3: Items involving \<eu> (and variants), \<ö>, or \<ü> in Jahn's (1927) data and their equivalents in other sources

	Jahn	Fonseca	Febres Cordero (1921)	Mirripú₂	Migurí	Tiguñió	Mucuchíes	El Morro Maripuyes	El Cenicero Mucuchíes
'six'	\<eu>	\<eu>	\<u>	\<u>	\<u>			\<i>	
'feet'	\<eu>	\<eu>, \<e>	\<ú>, \<uy>				\<u>		
'water'	\<eu>, \<u>	\<eu>, \<u>	\<u>	\<u>	\<u>		\<iú>, \<ué>		\<u>
'stone',	\<eu>		\<u>	\<u>					\<u>
'louse'	\<éü>	\<eu>	\<u>	\<u>					
'potato'	\<éu>	\<eu>, \<ú>, \<üi>	\<ú>, \<ui>, \<uí>	\<u>	\<u>		\<ui>	\<üi>	
'wind'	\<eu>, \<u>	\<u>	\<u>						

Table 4: Correspondence sets involving \<eu> (and variants) in Jahn's (1927) data

	Jahn	Fonseca	Febres Cordero (1921)	Mirripú₂	Migurí	Tiguñió		El Morro Maripuyes	El Cenicero Mucuchíes
'water'	\<ö>	\<o>	\<i>, \<u>	\<i>	\<i>				\<u>
'fire'	\<ö>	\<o>	\<u>	\<u>	\<u>			\<u>	\<u>
'cow, milk'	\<ö>	\<o>	\<o>						
'firewood, strength'	\<ö>, \<e>	\<e>, \<uo>		\<e>	\<e>			\<u>	
'hunger'	\<ö>					\<o>			

Table 5: Correspondence sets involving \<ö> or \<ü> in Jahn's (1927) data

5 Consonants

5.1 Prenasalization

Among the observations by Febres Cordero (1921:99), the following important statement is found:[3]

> Existía cierto sonido nasal, observado también por el viajero Wiener en el habla de algunas tribus del Amazonas, sonido que corresponde a algo así como débil ronquido, inarticulado; que se escapa más por la nariz que por la boca; y que no puede expresarse sino por una *n*, o más bien por una *m* inicial.

> A certain nasal sound, also observed by the traveler Wiener in the speech of some Amazonian tribes, existed, a sound which corresponds to something like a weak, inarticulated roaring, which escapes more through the nose than through the mouth, and which cannot be expressed other than by an initial *n*, or better by an initial *m*.

However, Febres Cordero (1921:142) explicitly notes this kind of pronunciation for only one of the items in his vocabulary, namely <umpu> ~ <mpu> 'sun'. This means that the special sound Febres Cordero describes cannot easily be identified in the data themselves.

Febres Cordero (1921:99-100) also indicates that the same type of sound exists in certain Guaraní words. Without giving a specific reference, he quotes an observation by a priest by the name of Zeledón on Chimila, a Chibchan language, according to whom an "*m*" is "melted" ("es licuante") with the "*r*" like in <mru> 'grain' and <mraamru> 'river current'. This sound may be a preglottalized tap /ʔɾ/ (cf. Quesada 2007:116). However, Chimila and Guaraní share a further distinctive phonetic trait that Febres Cordero more likely alludes to here, namely prenasalized stops. In Chimila, this is a redundant phonetic feature of voiced stops (Malone 2006:3; 2010:2), while Guaraní, the other language Febres Cordero mentions, features prenazalized obstruent phonemes /ᵐb/, /ⁿd/ and /ᵑg/ (Estigarribia 2017: 18, table 2.1). Other things being equal, prenasalization indeed seems to be a good interpretation of Febres Cordero's description of the sound, since it both accounts for the nasal character of the sound Febres Cordero describes, as well as for the presence of voicing which that can be inferred from Febres Cordero's "ronquido" (while *ronquido* in Spanish can refer to snorting, it is also commonly used e.g. for the roaring of deer). An alternative interpretation of Febres Cordero's description as referring to plosives with a nasal release is less plausible, since he states that the sound should be represented by initial <n> or <m>, which suggests a phonetic peculiarity which precedes the release of the plosive. Yet another possible interpretation, which would interpret the "ronquido" Febres Cordero mentions as aspiration, in line with the meaning 'snorting' which the word in Spanish also has, suffers from not accounting for the clear presence of nasal airflow which must be deduced from Febres Cordero's description.

Nevertheless, we must still assess whether there is any evidence of prenasalization in the actual data. This has to involve a distributional analysis of sounds and combinations of sounds that are candidate orthographic representations of prenasalized stops. These prominently include letters that –among other sounds– represent voiced stops in Spanish, i.e. <b>, <d>, and <g>. An assessment of these letters would also take into account the possibility that some transcribers missed possible prenasalization and/or assigned the associated voicing to the consonant itself. Also, nasal + stop combinations, i.e. <m> or <n> preceding the abovementioned symbols for voiced stops, must be evaluated.

The letter <b> occurs in the transcriptions of all authors mainly in two environments, namely intervocalically or preceded by <m>.

---

[3] Again, Febres Cordero (1921) does not specify what lect exactly he has in mind; the statement appears to be of a general nature.

As far as the intervocalic environment is concerned, some cases of variation in the transcriptions of a single author, e.g. <wihao> ~ <bijao> ~ <uihao> 'Heliconia' in Fonseca's (1955:291) Cuica data, strongly suggest that the sound transcribed here by <b> is not a stop but a fricative or approximant.

Spanish /b/ is realized as a stop [b] word-initially and following a nasal and as a homorganic fricative in all other environments. That intervocalic <b> in the Timote-Cuica transcriptions, alongside the unique example of the alternation in Fonseca's transcription above, represents a fricative is not only expected from the conventions of Spanish orthography, but is also shown by alternations of <b> with <v> and in one case also with <h> in Jahn's Cuica data. In addition, for the numeral 'nine', Jahn (1927) has <mabipita>, Fonseca (1955) <mabipita> ~ <mavipita> ~ <mavi pita> (erroneously giving the form as the numeral for 'eight' instead of 'nine'), and Febres Cordero (1921) <maipit>. Either a diachronic process of lenition along the lines of [b] > [β] > Ø was occurring in some lects. Alternatively, what is represented by Jahn and Fonseca as <b> or <v> and by Febres Cordero as Ø could be a sound that was easily be overheard and hence not represented at all, for which a bilabial fricative would be a very good candidate, too.

Febres Cordero (1921: 94) also offers a note that is related to this fricative sound:

> Verdad que no es fácil precisar si este sonido sea propiamente el suave de la labial b, o el más fuerte de la *v*, por la confusión que de ellas se hace al pronunciarlas. La *b* aparece convertida en *p* en *Mucumbís, Mucumpís*, nombres que indistintamente se dieron a una tribu y a un territorio."

> True, it is not easy to specify if this sound is the soft one of the labial b or the stronger one of the v, due to the confusion which is made of them when pronouncing them. The *b* appears converted into *p* in *Mucumbís*, *Mucumpís*, names which were indistinctly given to a tribe as well as a territory.

Having established that intervocalic <b> in a significant number of cases in which it occurs intervocalically could represent a fricative (whether [β] or [v]), one is, with the exceptions noted above, left with a remarkable distribution of <b>. Otherwise it occurs almost exclusively in the digraph <mb>, as in the placename Febres Cordero mentions. Under the assumption that this digraph indeed represents the hypothesized prenasalized [ᵐb], some of the exceptions become explainable, and I believe, turn out to be not inconsistencies but rather additional evidence for [ᵐb].

One particularly frequent alternation is that with <nb>, as in Jahn's Cuica form <kuanbéuch> ~ <kambéuch> 'staff, cane'. These uncertainties as to the place of articulation of the nasal are explainable if its duration, as expected in prenasalization, would be relatively short. However, they are also consistent with Arrieta's (1998) interpretation as indicating the nasal quality of the preceding quality. Yet other evidence points clearly towards the interpretation in terms of prenasalization. Notable is the pair <beuch> ~ <timbeuch> 'wood' in Fonseca's Cuica data. <timbeuch> features the plural marker <ti>. In the word-medial environment created by the presence of this prefix, the sound transcribed word-initially as <b> becomes <mb>. Since Spanish phonotactics do not allow for nasal-consonant clusters in word-initial position, it is expectable that native speakers of Spanish were induced to not transcribing the prenasalization in this position with a separate symbol, but simply as <b>. However, the quality of the sound becomes visible with added morphology. That prenasalized stops did occur in initial position in Cuica is shown rather clearly by another instructive set of alternations in Jahn's data, namely <a be> ~ <be> ~ <abé> ~ <mabe> 'where?' The form with the highest token frequency is <be>. That the <b> in this form represents a prenasalized stop can be seen from the variant forms with an initial <a>, in which case the voiced and continuant features of the prenasalization must have been heard as vocalic; that there was a nasal can be seen from the variant <mabe>. The representation of the sound in initial position is thus parallel to Febres Cordero's <umpu> ~ <mpu> 'sun', of which we know explicitly that the sound in question was present. The development of the spelling of placenames is also instructive: The place today called Maruchí is attested as <ibaruchí> in the 17th century and as <buruchí> in the 16th century (Salas 1908: 179). Fonseca (2005:274) mentions some interesting variations in the spelling of rivers in colonial documents which are characterized by an alterna-

tion of <m> and <b>: <bomboy> ~ <momboy>, <bimbate> ~ <mimbate>, <bitimbis> ~ <mitimbis>, <pocó> ~ <bocó>, <bonay> ~ <monay>. Such alternations, which show phenomena that are entirely parallel to those observed in the sets <a be> ~ <be> ~ <abé> ~ <mabe> and <umpu> ~ <mpu>, fit seamlessly into the chain of evidence in favor of the presence of prenasalized stops.

The letter <d>, Febres Cordero (1921:96) says, does not exist safe for some rare exceptions. Indeed, <d> is very rare in the extant data. All transcribers use the letter virtually only in the combination <nd>; exceptions are mostly loanwords or have the combination <md> instead of <nd>. Analogously to the interpretation of <mb>, <nd> is interpretable as a prenasalized plosive [^{n}d] to account for the striking co-occurrence tendencies. The identification of this sound in intial position can also be done analogously to that of [^{m}b]. Jahn's Cuica form <dú> ~ <du> ~ <a dú> and Fonseca's <du> ~ <udu> 'with, together' are consistent with the presence of [^{n}d] in word-initial position. Notable is also that the authors in this case use <u> instead of the <a> employed in the case of [^{m}b], that is a vowel with a point of articulation further back for a nasal with a point of articulation further back (though it is also significant and probably influential that the vowel following the prenasalized stop is also back in the relevant forms).

<nt> is a possible alternative transcription that is relatively frequent. In several instances, Jahn and Fonseca's <nd> in Cuica correspond to <nt> in Febres Cordero and Lares's Timote data, which is compatible with either different transcription habits or indeed a phonetic difference between Timote and Cuica.

The distribution of <g> in the extant data on Timote-Cuica is the most difficult to make sense of, not least because the letter has a number of functions in Spanish orthography depending on the environment.

In the Timote-Cuica data, <g> occurs frequently next to an <n>. In this context it is not uncommon either that odd vowel symbol sequences such as <ue>, <u>, <üé> etc. are found in the vicinity. Also, unlike <b> and <d>, <g> occurs followed by <n>. Examples of these phenomena in Cuica, all from Jahn (1927), are <ki-köngk> 'nose' <tit-kushpapeúgn> 'bones', <kiu-guegn> 'flour', <kiu-kiangüe> 'millstone', <kiu-kuagüé> 'k.o. locust', <ki-fungt> 'scorpion', <kiu-chaséugn> 'prickly pear'. <nk> clusters are also found.

It is quite likely that <ng> and perhaps also <gn> and <nk> in a subset of instances represents the velar nasal [ŋ], an idea strongly supported by the transcription habits of colonial grammarians and lexicographers of the Andes (Urban submitted). If this were indeed the case, the sequence <ngk> in <ki-köngk> 'nose' could plausible represent a prenasalized velar stop [^{ŋ}g] (cf. Fonseca <kug>), but clearly, demonstrating the existence of such a sound at the velar position is the most difficult of all places of articulation.

There is no clear evidence for the prenasalized stops to be in complementary distribution with the voiceless stops, which indicates that phonologically they cannot easily be treated as allophones of the same phonemes. In particular, the fact that they occur in initial position as clearly indicated by Febres Cordero himself and as also demonstrated by the data strongly suggest a phonemic function. For the bilabial and alveolar position, even apparent minimal pairs can be found in Jahn's Cuica data: <shambú> 'tomorrow' vs. <shapú> 'to put' and <shandú> 'ignite' vs. <shatú> 'to tell', a verb that can be extracted from several phrases in Jahn's data.

This analysis is fundamentally different from that of Arrieta, who posited a single series of basically voiceless stops which have voiced allophones following nasal vowels (according to her, as we have seen indicated by syllable-final <m> and <n>). Arrieta (1998:93) did note the alternation <du> ~ <udu> 'with' as one of the exceptions to the observation that <d> is regularly preceded by <n>. This is left unexplained in her account, but finds a perfectly plausible explanation assuming, as proposed here, that what is represented is a prenasalized stop in initial position. Also left unexplained is the presence of <nt> sequences, since in her account the voiced allophone would be expected here.

In spite of the strong evidence for prenasalized stops in the data, these remain difficult to distinguish from nasal vowels which indeed most likely occurred too and were, as correctly analyzed by Arrieta, also indicated in some instances by <m> and <n>.

5.2 The "Strong" <c>

Febres Cordero has more observations on the sound of Timote speech as heard by him. He (1921:94 -95) mentions that

> [l]a *c* tenía un sonido muy fuerte en las sílabas *ca, co, cu* y al final de las voces. Algunos gramáticos, entre ellos el erudito don Julio Calcaño, han empleado al efecto la *k* en vez de la *c*, y en antiguos manuscritos hemos visto usada la *q* combinada directamente con *a* y *o*, en *Moqojún* y *Moqaquetá*, por ejemplo, acaso para reforzar la pronunciación. La *c*, al final de algunas voces, aparece convertida en *que*, como lo hemos observado en *Estictec, Misisic, Moconoc*, escritos también *Esticteque, Misisique, Moconoque* ; y se observa lo mismo en *Micbají*, escrito *Miquibají*.

> The *c* had a very strong sound in the syllables *ca*, *co*, *cu* and at the end of words. Some grammarians, among them the erudite Julio Calcaño, appropriately employed the *k* instead of the *c*, and in old manuscripts we have seen the *q* combined directly with *a* and *o*, for instance in *Moqujún* and *Moqaquetá*, perhaps to reinforce the pronounciation. The *c*, at the end of some words, appears changed into *que*, as we have observed in *Estictec*, Misisic, Moconoc, which are also written *Esticteque*, *Misisique*, *Moconoque*; and the same is observed in *Micbají*, written Miquibají.

The characterization of the sound as particularly "strong" may be interpreted in many different ways. It could allude to a uvular point of articulation, to aspiration or velarization ([kʰ] or [kˣ]), or to glottalization. However, the solid evidence that would be required to postulate either is lacking. A thorough search of the available data would be needed to identify patterns that would allow to distinguish [k] from any of the other hypothetical pronunciation. Since I have not found any up to this point, I have no good evidence to assume that <k> represents anything other than a velar voiceless stop [k].

5.3 Fricatives and Affricates

A further aspect of pronunciation that Febres Cordero (1921:95-96) chose to comment on are fricatives and affricates:

> La *ch* usada en muchísimas voces indígenas, no corresponde ciertamente al sonido fuerte que se le da en castellano, sino a uno mucho más suave como en el nahuatl y en el goajiro, casi equivalente a la simple *s*, según puede comprobarse con muchos ejemplos : *chep, quichán*, que se pronuncian *sep, quisán*, silvando [sic!] la *s* más de lo regular. Don José Ignacio Lares dice a este respecto, que la *ch* de los indios tiene el sonido de la *sh* inglesa o *sch* alemana.

> The *ch* used in very many indigenous words does not definitely correspond to the strong sound which it is given in Spanish, but to a much softer one like in Nahuatl or Goajiro, almost equivalent to the simple s, in accordance to what one can verify with many examples: *chep*, *quichán*, which are pronounced *sep*, *quisán*, the *s* hissing more than normal. José Ignacio Lares says in this respect that the *ch* of the Indians has the sound of the English *sh* or the German *sch*.

Indeed, Ernst (1885:194fn1) relates a note from Lares that <ch> in his Timote data should always be pronounced like the English <sh>, i.e. as a postalveolar voiceless fricative [ʃ]. Fonseca (2005:268), too, mentions that <sh> in his Cuica data represents the "s gruesa", i.e. [ʃ]. The presence of some instances of <sch> in Jahn's Cuica transcription, which is the orthographic representation of [ʃ] in German, suggests the presence of this sound as well, even though it is confusing to note that occasionally <sch> makes an appearance in Fonseca's data as well, sometimes even where Jahn has <sh>. Nevertheless, from the clarity of the description and from available sources on Guajiro and (classical) Nahuatl, the languages which are mentioned alongside English and German for comparison, there is no indication that Febres Cordero could have had in mind a fricative other than [ʃ] such as, for instance, an alveolo-palatal [ɕ] (Mansen and Captain 2000:796; Andrews 2003:27).

Febres Cordero (1921:101) has more to say on what appear to be affricates or affricate-like sounds:

> Como en otras lenguas americanas, existía una consonante que indistintamente se ha traducido en castellano por *t*, por *z*, por *s,* y hasta por *ch*, por corresponder al sonido *tz*. *Tzirup* sería, por ejemplo, la ortografía más propia de *Sirup*, lumbre, candela encendida, como la hemos escrito para mayor claridad. Creemos que esta consonante indígena sea la que ocurre en la citada voz *Caquetía*, escrita también *Caquesia* y Caquexia ¿ No será *Caquetzia* ?

> Like in other American languages, a consonant which has been indiscriminatingly represented in Spanish by *t*, by *z*, by *s*, and even by *ch*, for corresponding to the sound *tz*. *Tzirup*, for instance, would be the more appropriate orthography of *sirup* 'light, lit candle', which is how we have written it for greater clarity. We believe that this indigenous consonant is that which occurs in the mentioned word Caquetía, which is also written Caquesia and Caquexia. Could it not be Caquetzia ?

The letter combination <tz> does not normally occur in the spelling of words from Spanish origin but does occur in loanwords from Basque as well as a number of loanwords from Amerindian languages. It is pronounced [ts].

Arrieta (1998:94) assumes that Timote-Cuica featured just one fricative phoneme with a wide range of realizations. This claim is based on alternations like the following (in her orthography):

(2) a.   *kizán ~ kishám ~ hushám ~ kichan ~ husham*   'head'
    b.   *tzirúp ~ sirup ~ chirup*   'fire'
    c.   *ti-marsot ~ ti-marzot*   'whip'
    d.   *ti-tisán ~ ti-ti-san ~ ti-ti-sán ~ ti-ti-zan*   'flies'
    e.   *shikás ~ chikás*   'bell pepper'
    f.   *timbó ~ timbós ~ timpósh ~ timposh*   'tuber'
    g.   *chés ~ shes*   'God'

However, each author's transcriptions reveal some peculiarities. Jahn, for instance, not ever uses <z> for Cuica. That mere transcription habits rather than allophonic variation are at play is shown by systematically assessing the correspondences sets of items which Jahn transcribes with <ts> with those of Fonseca and, where possible, also other authors. Table 6 shows those instances where Jahn's <ts> corresponds to a letter or letter sequence other than <ts> (cases where <ts> corresponds to <ts>, which also exist, are analytically vacuous).

The sets in table 6 yield the correspondences in table 7. Clearly identifiable extra morphology (e.g. the article *kiu*) have not been taken into account in setting up the correspondences. That notwithstanding, there is not always absolute clarity on the proper interpretation. For instance, should the equivalent to <ts> in Jahn's and Fonseca's Cuica forms in Febres Cordero's <cacsún> be considered <cs> or only <s>?

	Jahn	Fonseca	Febres Cordero (1921)	Mirripú[1]	Mirripú[2]	El Morro Maripuyes
'child'	\<kiu-keúnts\>, \<kiu keunts\>	\<kiú keuntz\>, \<cae-unts\>, \<keuntz\>, \<cae-untz\>, \<kiu keums\>, \<eunts\>				
'páramo'	\<kiu-gánk\>, \<tsugánk\>, \<pshugank\>, \<tsu gank\>	\<gane\>, \<kiu gank\>, \<tsu gank\>				
'six'	\<katséun\>	\<katseunt\>, \<katseun\>	\<cacsún\>		\<kasum\>	\<kapsín\>
'It's still early'	\<e wo tsi tsuír\>	\<who shi tskui\>				
'evening'	\<wotsúi\>, \<wok-tsuí\>	\<wksui\>, \<woktskui\>, \<woktsui\>, \<woksui\>				
'fleas'	\<ti-kits\>	\<tit ikit\>, \<titcits\>,\<tit ikits\>, \<titikits\>, \<ti tikts\>, \<ti tikits\>	\<quis\>		\<kis\>, \<quis\>	
'corn cob'	\<kiú-tsaós\>	\<ih kiu tsaos\>, \<ti tsaos\>, \<kiu tsaos\>, \<tzao\>				
'mountain strawberry'	\<kiu-tsatséu\>	\<tzatzeu\>, \<tzatz\>				
'small, a little'	\<sits\>, \<dits\>	\<sits\>, \<shic\>, \<kashists\>, \<kasits\>				
'exclamation of sweet taste'		\<tsa\> \<tza\>, \<tha\>, \<tja\>				
'exclamation of sour taste'		\<he\>, \<tze\>, \<tse\>				

Table 6: Correspondence sets for items transcribed with \<ts\> in Jahn's (1927) data. Only heterogeneous correspondence sets are shown.

	Jahn	Fonseca	Febres Cordero (1921)	Mirripú₁	Mirripú₂	El Morro Maripuyes
'child'	<ts>	<ts>, <tz>, <ms>				
'páramo'	<ts>, <psh>	<ts>				
'six'	<ts>	<ts>	<ks>		<s>	<ps>
'It's still early'	<ts>	<ts>				
'evening'	<ts>	<ts>, <s>				
'fleas'	<ts>	<ts>, <t>	<s>		<s>	
'corn cob'	<ts>	<ts>, <tz>				
'mountain strawberry'	<ts>	<tz>				
'small, a little'	<ts>	<ts>, <k>				
'exclamation of sweet taste'		<ts>, <tz>, <h>, <j>				
'exclamation of sour taste'		<ts>, <tz>, <h>				

Table 7: Heterogeneous correspondences for items transcribed with <ts> in Jahn's (1927) data

In spite of such issues, the correspondences in table 6 show at least two things: first, as already alluded to, there appears to be a significant component of merely orthographic variation especially when comparing Jahn's with Fonseca's Cuica data. Fonseca has a certain preference for <tz> over <ts>, but since <tz> and <ts> alternate in transcriptions of the very same item in his data, it is not unreasonable to suppose (but neither possible to exclude) that phonetic variation is implied. Second, the variation for items involving the digraph <ts> in Jahn's data is not as unconstrained as Arrieta's (1998) statement would suggest. For one, there is not a single instance where <ts> would alternate with letter combinations that would suggest a palatal fricative or affricate such as <ch>, <sh>, or <sch>. Rather, without wishing to gloss over the paucity of data and the fact that transcription like <ks> and <ps> suggest a rather complex phonetic reality, the data point to a reasonably regular correspondence between <ts>/<tz> in the Cuica data of Jahn and Fonseca with <s> in the Timote data of Lares and Febres Cordero. This may well reflect a correspondence between an alveolar affricate [ts] to a homorganic non-affricate [s]. From the point of view of phonology, the facts suggest the possibility that the sounds behind <ts> ~ <tz> ~ <s> must be distinguished from those behind <ch>, <sh>, and <sch>, a possibility that would, however, require a more detailed investigation.

5.4 Velar and Glottal Sounds

Finally, Febres Cordero (1921:97) offers an informative statement relating to velar and glottal sounds. He says:

> El sonido fuerte de la *j* tampoco es indígena. En las voces en que aparece hoy esta letra, a la verdad abundantes en los Andes, pudiera emplearse con más propiedad fonética la *h*; porque se trata de señalar una breve pausa o aspiración. Así lo hicieron los primeros españoles, escribiendo v. g., *Buhay, Tinoho*, convertidos después en *Bujay* y *Tinojo*. También solían representarla por *s*, como en *hisjut* o *hisut*, y por contraction *sut*, que significa tres.

> The strong sound of *j* is also not indigenous. In the words in which this letter appears today one could with more phonetic adequacy use *h*, because it is a matter of indicating a short pause or aspiration. Thus did the first Spaniards, writing for instance *Buhay*, *Tinoho*, which were later converted into *Bujay* and *Tinojo*. They also used to represent the sound by *s*, as in *hisjut* or *hisut*, and by contraction *sut*, which means three.

The "stong sound of *j*" that Febres Cordero mentions must be the velar fricative [x]. At any rate, that [x] is not among the sounds of Cuica can be confirmed by the evidence from loanwords, where initial [x] in *Juán, Julián, and Jueves* is replaced by [ʃ], presumably the fricative sound at the point of articulation closest to the velar position available in Timote-Cuica; the respective forms are transcribed as <Shuán> ~ <Shjuán> (Jahn), <shuan> (Fonseca), <shukian> (Fonseca) and <shueves> (Fonseca). The characterization of the sound as a "pause" on behalf of Febres Cordero speaks more in favor of an interpretation as a stop sound [ʔ], while "aspiration" points towards [h].

Similarly, Spanish [ɣ] and/or [ɰ] are replaced by other sounds in the adaptation to Timote-Cuica. The replacing sound is also [ʃ] in <chaguarden> ~ <chaguardén>, a loanword in the Timote variety of Mirripú₂ from Spanish *aguardiente*. The replacement is accompanied here by aphaeresis of the initial unstressed vowel. Replacement of [ɣ] / [ɰ] medially is somewhat less systematic. [k] appears to be the substituting sound in the various transcriptions of the form for 'wheat', borrowed from Spanish *trigo*. The transcription <tircüé> and <tirkué> occurring in Febres Cordero's El Morro Maripuyes data suggest that the original continuant character of the original Spanish sound may have been preserved, at least in this variety, by a glide. Spanish *enagua(s)* 'skirt' comes out as <najuá> in El Morro Maripuyes, suggesting a glottal point of articu-

lation in line with Febres Cordero's comments on <j> made above.[4] Such substitution patterns have already been observed by Fonseca (2005:268).

Febres Cordero suggests that the glottal sound was initially represented by <h>. There is also other very strong evidence for a glottal sound –whether it was [h], [ʔ], or both in variation with one another– in the Cuica data: Jahn has a few forms in which the same vowel letters are repeated adjacent to one another, as in <aá> 'to complain'. That this is not representing a long vowel is shown by the corresponding transcription in Fonseca, who has <a a>. Moreover, Fonseca (1955:303) explicitly states that <kuu> 'to sleep' is disyllabic, which is incompatible with an interpretation as a long vowel. Rather, there must be two vowels separated by a consonantal sound that has not been transcribed; this particular piece of evidence, is, however, complicated by the fact that there are the variant transcriptions of the verb to 'sleep' which are <keum> and <kun>. Nevertheless, such forms are the clearest evidence for the presence of a glottal sound in the available data. It is likely to have been present elsewhere, too. For instance, Arrieta (1998:93) posits a velar fricative [x] as an allophone of /k/ on the basis of alternations as the following (in her orthography):

(3) a.   *xok ~ köng ~ kong*   'nose'
    b.   *numbuk ~ numbuh ~ numbúh*   'salt'
    c.   *chingkang ~ chingkanx*   'flour'
    d.   *sukhós ~ suhxos*   'trunk'

An alternative interpretation in the light of Febres Cordero's explicit statement that the "strong sound of *j*", i.e. the velar fricative, is absent in Timote-Cuica, debuccalization of /k/ is an alternative account for the above alternations.

6 Conclusion

In this article, I have intended to incorporate first-hand observer's statements into the reconstitution process of the Timote-Cuica languages of the Venezuelan Andes. These comments have hitherto not been taken into account in interpreting the data. Such statements are not available for all sounds of the languages and hence their interpretation cannot yield a complete reconstitution. Nevertheless, the result of this initial investigation partially confirms earlier interpretations, but also suggest that the picture of Timote-Cuica phonetics and phonology might need rethinking in other regards. As far as the vowel system is concerned, observer's statements actually reinforce Arrieta's (1998) reconstitution of phonetic and phonological structure in that they, too, point towards the existence of nasal vowels and a high-central vowel. As far as the consonant system is concerned, some hitherto unrecognized features could be suggested. In the first place, observer's perceptions on certain sounds that existed in the languages and the subsequent analysis of the data themselves suggest that there were two plosive series, plain and prenasalized. Likewise, for affricates and fricatives the same manner to proceed suggests preliminarily that rather than a single fricative phoneme with a wide range of realizations, fricatives/affricates at at least two places of articulation, alveolar and palatal, might need to be distinguished. The data also show evidence for a glottal on the phonetic level.

On the methodological level, the investigation adds authority to Constenla Umaña's (2000) recommendation that in the reconstitution of phonetic and phonological structures in premodern and imperfect language data it is not only the data themselves matter, but also all ancillary information provided by those who had some exposure to the languages as they were actually spoken. It also underscores that data must be treated in a philologically adequate manner. Some of the insights of the present article could, in spite of the complex interrelations of the available data, only be gained by separating as strictly as possible data of different transcribers instead of lumping them into one big corpus of data. This is not only adequate because

---

[4] Note that Febres Cordero transcribes the initial sound in the adapted version of Spanish *calzón* 'pants' also with <j> in El Morro Maripuyes. Why [k] should have been substituted here is unclear.

variation at the dialectal level (Dench 1995) or even above is a real possibility, but also because this separation is crucial for the identification of covert patterns in the data, the task which reconstitution is all about.

7 Acknowledgments

8 References

Adelaar, Willem F. H., with the Collaboration of Pieter C. Muysken. 2004. *The languages of the Andes* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Andrews, J. Richard. 2003. *Introduction to Classical Nahuatl*. Revised edn. Norman: University of Oklahoma Press.

Arrieta E., Anita. 1993. Tipología morfosintáctica del Timote. *Filología y Lingüística* 19(2). 99-110.

Arrieta, Anita. 1998. Tipología fonológica del Timote. *Filología y Lingüística* 24(1). 85-100.

Briceño-Iragorry, Mario. 1929. Procedencia y cultura de los Timoto-Cuycas. *Anales de la Universidad Central de Venezuela* 17 (2). 156-183.

Broadbent, Sylvia M. 1957. Rumsen I: methods of reconstitution. *International Journal of American Linguistics* 23(4). 275-280.

Calcaño, D. Julio. 1886. *Resumen de las actas de la Academia Venezolana correspondiente de la Real Academia Española*. Caracas: Imprenta Nacionál

Chamberlain, Alexandre F. 1910. Sur quelques familles linguistiques peu connues ou presques inconnues de l'Amérique du Sud. Étude d'orientation linguistique. *Journal de la Société des Américanistes* 7. 179-202.

Constenla Umaña, Adolfo. 2000. La restitución: un método lingüístico reconstructivo sincrónico. *Filología y Lingüística* 27(2). 161-180.

Cysouw, Michael & Jeff Good. 2013. Languoid, doculect and glossonym: formalizing the notion 'language'. *Language Documentation and Conservation* 7. 331-359.

De Santiago, Pedro A. 1956. *Biografías trujillanas. Homenaje a Trujillo en el cuarto centenario de su fundación, 1557-1957*, vol. 1. Caracas: Ediciones Edime.

Dench, Alan. 1995. Comparative reconstruction. In John Charles Smith & Delia Bentley (eds.), *Historical Linguistics 1995. Selected papers from the 12th international conference on historical linguistics, Manchester, August 1995, vol. 1: general issues and non-germanic languages* (Current Issues in Linguistic Theory 161), 57-72. Amsterdam/Philadelphia: John Benjamins.

Ernst, A. 1885. Ueber die Reste der Ureinwohner in den Gebirgen von Mérida. *Zeitschrift für Ethnologie* 17. 190-197.

Ernst, A. 1886. Sprachproben der Ureinwohner von Venezuela. *Zeitschrift für Ethnologie* 18. 372-373.

Estigarribia, Bruno. 2017. Grammar sketch of Paraguayan Guarani. In Bruno Estigarribia & Justin Pinto (eds.), *Guarani linguistics in the 21st century* (Brill's Studies in the Indigenous Languages of the Americas 14), 7-85. Leiden/Boston: Brill.

Febres Cordero, Tulio. 1921. *Historia de los Andes. Procedencia y lengua de los Aborigenes*. Mérida: Tip. "El Lápiz."

Fonseca, Amílcar. 2005. *Orígenes Trujillanos*. Trujillo: Fondo Editorial Arturo Cardozo.

Goddard, Ives. 1973. Philological approaches to the study of North American Indian languages: doc-

uments and documentations. In Thomas A. Sebeok (ed.), *Current trends in linguistics, vol. 10: linguistics in North America*, 727-745. The Hague/Paris: Mouton.

Jahn, Alfredo. 1927. *Los aborígenes del occidente de Venezuela. Su historia, etnografía y afinidades lingüísticas*. Caracas: Lit. y Tip. del Comercio.

Lares, José Ignacio. 1907. *Etnografía del Estado Mérida*. 2nd corrected and augmented edn. Mérida: Imprenta del Estado.

Lares, José Ignacio. 1918. Andes Venezolanos: vocabularios. *De Re Indica* 1. 35-36.

Malone, Terry. 2006. Tone and syllable structure in Chimila. *International Journal of American Linguistics* 72(1). 1-58.

Malone, Terry. 2010. Nasality and nasal prosody in Chimila. *International Journal of American Linguistics* 76(1). 1-41.

Mansen, Richard & David Captain. 2000. El idioma wayuu (o Guajiro). In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia. Una visión descriptiva,* 795-810. Santafé de Bogotá: Instituto Caro y Cuervo.

Pérez Hernández, Francisco Javier. 1988. *Historia de la lingüística en Venezuela desde 1782 hasta 1929*. San Cristobal: Universidad Católica del Táchira.

Quesada, J. Diego. 2007. *The Chibchan languages*. Cartago: Editorial Tecnológica de Costa Rica.

Rivet, P. 1927. La famille linguistique Timote (Venezuela). *International Journal of American Linguistics* 4(2/4). 137-167.

Salas, Julio C. 1908. *Tierra-firme (Venezuela y Colombia). Estudios sobre etnología e historia*. Mérida: Tip. de "Paz y Trabajo."

Urban, Matthias. 2015. Notas sobre el sistema de numeración de las lenguas Timote-Cuica. *Boletín Antropológico* 33. 53-69.

Urban, Matthias. Submitted. The representation of the velar nasal in colonial grammars and other premodern sources on the languages of the Central Andean region. In *Proceedings of the Revitalizing older linguistic documentation (ROLD) colloquium 'From Greenland to Patagonia'*.