

# **Cholón and the linguistic prehistory of northern Peru: triangulating toponymy, substrate lexis, and areal typology\***

Matthias Urban  
University of Tübingen

*At the eve of Spanish conquest, northern Peru is thought to have been home to a multitude of languages of relatively modest geographical extension, especially when compared with the widespread Quechuan and Aymaran languages. In this contribution, I suggest the possibility that a language or several languages relatively closely related to Cholon were spoken in a much wider part of northern Peru than that in which Cholon is historically attested. A prior “Cholonoid” area might have covered not only the western part of today’s San Martín department, but also almost the entire department of Cajamarca as well as parts of La Libertad and Amazonas. This interpretation results from a triangulation of three independent lines of evidence, namely the toponymic record, substrate lexis in the local variety of Quechua at Chachapoyas, and typological properties of the extinct northern Peruvian languages.*

## **1. Introduction**

The original linguistic diversity of the northern Peruvian Andes is scarcely visible today. In the course of the centuries, Spanish has replaced most of the individual languages once spoken on the coast, in the highlands, and on the eastern slopes of the Andes. Today, what remains are three different Quechua varieties in the areas of Ferreñafe, Cajamarca, and Chachapoyas, the latter in particular also threatened by extinction. Yet, studies commencing in the last decades of the 20<sup>th</sup> century have succeeded in reconstituting a reasonable approximation of the complex linguistic landscape that existed in the past (e.g. Torero 1986, 1989, 1993, 2002, Adelaar 1988, Cerrón-Palomino 2004, Urban 2019b).

On the coast, which in northern Peru forms a narrow strip of highly arid land that quickly gives way in the east to the Andean highlands, at least four distinguishable languages were spoken at the point of European contact.<sup>1</sup> Moving from north to south, these languages are conventionally called Tallán, Sechura, Mochica, and Quingnam. Tallán and Sechura were languages of the coast of the Piura department, the former spoken in at least the settlements of Colán and Catacaos, the latter in the town of Sechura and surroundings. What remains of these languages are short wordlists, place- and personal names, and some vocabulary items related to the local culture in Spanish (Urban 2019b). Mochica, once spoken in the coastal areas of Lambayeque, the northern part of the La Libertad region, and probably in the valley of the upper Piura river, is the best documented of the coastal languages. It also survived longest, until the beginning of the 20<sup>th</sup> century. Personal names from 16<sup>th</sup> century Cajamarca suggest that there were at least pockets of people of Mochica origin in the highlands, too (cf.

---

\* I thank Willem F.H. Adelaar, Rodolfo Cerrón-Palomino, Nicholas Q. Emlen, Matthias Pache and two anonymous referees for commenting on this paper, although they do not necessarily share the views expressed in this article and are not responsible for errors of fact or interpretation. I also thank Arjan Mossel for help with GIS, for who the same disclaimer applies. Work on this article was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 295918 (“The Linguistic Past of Mesoamerica and the Andes: A search for early migratory relations between North and South America”, PI: Willem F. H. Adelaar) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project No. UR 310/1-1 (“The language dynamics of the Ancient Central Andes”, PI: Matthias Urban).

<sup>1</sup> Then as today, the coastal population was concentrated in a series of fertile river valleys which dissect the arid lands longitudinally in fairly regular intervals. People associated with the coast culturally and linguistically probably populated also parts of the upper river valleys.

Rostworowski de Diez Canseco 1985 and Urban 2019b for review of other evidence). Still further south, the people of the coast of northern Peru spoke a language known as Quingnam. In terms of both level of documentation and date of extinction it is the opposite of Mochica, having become extinct very early with only the most minimal documentation available. The southern limit of the Quingnam-speaking zone is still poorly defined (cf. discussion in Salas García 2010 and Urban 2019b). In spite of the scarce documentation of the coastal languages other than Mochica, the available lexical record shows some similarities that can be attributed to relatively intensive language contact (Urban 2019b).

In the highlands to the east of the coastal desert plain, there is a stronger presence of Quechuan than in the coastal areas. In the Lambayecan province of Ferreñafe and in the highlands of Cajamarca, different varieties are spoken still today. These areas are best thought of as the surviving pieces of a complex linguistic mosaic that also involved non-Quechuan languages. One of these was the Culli language, which shares the fate of minimal documentation with Tallán and Sechura. Toponymy (Adelaar 1988, Torero 1989) suggests that Culli was once spoken in Southern Cajamarca (Cajabamba province), the highlands of La Libertad (Otuzco, Sánchez Carrión, and Santiago de Chuco provinces), and the north of Ancash (Pallasca province). However, toponymy also suggests that the linguistic landscape in the northern Peruvian highlands was once still more complex, in particular in Cajamarca. Torero (1989) has identified two toponymic areas that are interpreted as the remnants of local languages once spoken there. These areas are defined by the characteristic endings -den (with variants -don, -ten, -ton, -din, -tin; the ending is stressbearing with great frequency) and -cat (with variants -cot, -gat, and -got and -cate, -cote, gate, -gote) respectively.<sup>2</sup> The voiced variant of the ending -cat very frequently occurs after nasals (Valqui Cauqui 2004), as is common in the Central Andes. The Den area, according to Torero (1989: 230), is exclusively Cajamarcan, extending across the entire provinces of Contumazá, San Miguel, Hualgáyoc, and Santa Cruz as well as parts of Cajamarca, Chota, Cutervo, and Celendín. Contumazá is a center of density.<sup>3</sup> Again according to Torero (1989: 232), toponyms which feature a high vowel in the ending (-din, -tin) occur typically on the margins of the area. There are some hybrid Quechua-Den toponyms, and possibly also two cases of hybrid Spanish-Den toponyms (Andrade Ciudad 2010: 174). In figure 1, which uses public domain data from the GEOnet Names Server of the US National Geospatial Intelligence Agency, Den area toponyms are plotted using a triangle symbol (placenames that fit by their shape, but that are known to have an etymology that is irrelevant for present purposes, were removed manually before plotting; examples include El Jardín, Edén, Wadington, etc.). The Cat area as identified by Torero (1989: 234) covers the entire Cajamarca department, too, but also extends to the highland provinces of the La Libertad department, the highlands of Lambayeque's Ferreñafe province, and the Utcubamba, Bagua, Luya, and Chachapoyas provinces of the Amazonas department, with isolated cases in the highlands of Piura and near the coast. Cat has a very high density of occurrence in Cajamarca, too. The locations of Cat toponyms are indicated in figure 1 by circles (also here, placenames with clear etymologies which make them irrelevant for present purposes have been removed manually).<sup>4</sup> Thus, in significant parts of their respective areas

<sup>2</sup> I use “ending” here as a technical term for a (set of) recurrent final sequences in toponyms that cluster in geographical space and suggest a common origin. The morphosyntactic status of the “ending” in the language from which it stems is not given; it could be an inflectional or derivation suffix or a full noun. The term “ending”, therefore, is not meant to suggest a suffix necessarily.

<sup>3</sup> In the following, I will use “Den” and “Cat” as labels for the respective toponymic areas and the hypothetical language(s) that are responsible for their formation. When referring to a variant of the actual ending that defines them, I keep on using the hyphenated form, e.g. -den or -cat.

<sup>4</sup> Torero (1989: 237) also identifies a set of toponyms with final -can, -gan, -con, or -gon that overlaps in its distribution with the Cat area. This resembles the Culli word for ‘water’ which is also frequent as a toponymic ending. However, Torero (1989: 237) appears to prefer to keep the two toponymic sets separated analytically.

Cat and Den toponyms overlap geographically; in the south, Cat also overlaps with the Culli area (not plotted in figure 1). Given the existence of hybrid Quechua-Den toponyms (Andrade Ciudad 2010: 174), its presence must overlap temporally with the presence of Quechua in northern Peru. The relation of the respective languages in geographical space and through time accordingly has required the postulation of sophisticated theories (Andrade Ciudad 2010).

Moving further to the east, where the Andes become lower again to finally give way to the western Amazonian lowlands, again a different linguistic picture obtains. The steep valley of the Marañón river can be conceived of as a frontier between highlands and eastern slopes which is also linguistically relevant. In the area around Jaén, already to the east of the Marañón valley, colonial Spanish reports mention as many as eight distinct languages in a very confined geographic space. The same reports cite between three and five words only for each. Torero (1993) managed to tentatively suggest affiliations with Amazonian languages for some of the Jaén languages on the basis of just this information; other languages remain unaffiliated. Further south, Chachapoyas Quechua is still spoken by a relatively small number of elderly people in the Chachapoyas province of the Amazonas department.

There is strong evidence from toponymy, personal names, and ethnohistory to suggest that Quechuan replaced an undocumented non-Quechuan language, conventionally called Chacha, in the Chachapoyas area relatively late in prehistory or early in prehistory (Taylor 1990). Typical Chachapoyas toponyms ending in -mal, -lon-, and -lap (Torero 1989: 238) cannot be reconciled with a Quechuan origin. Alongside -cat, which also extends to Chachapoyas, other characteristic endings are -huala (Taylor 1990) and -oc or -ox (Valqui Culqui 2004). Figure 1 shows the clustering of toponyms in -mal, lon, and -lap, plotted as pentagons, rectangles, and crosses respectively, in the Chachapoyas region as well as the extension of the Cat area to Chachapoyas. Also, most of the personal names from the Chachapoyas region (assembled in Zevallos Quiñones 1966 and Rivarola 2004) are decidedly un-Quechuan. Interference from the original language of Chachapoyas may also be responsible for some drastic changes in Chachapoyas Quechua, such as shift of stress to the initial syllable and subsequent reduction of unstressed vowels (Taylor 1979). To the south of Chachapoyas, Hibito and Cholón were the dominant local languages on the eastern slopes of the Northern Andes on the interfluvium of the Marañón and Huallaga valleys from roughly Juanjuí in the north to Tingo María in the south (Alexander Bakkerus 2005: 33). Hibito and Cholón will play a crucial role in the present article, so that some details on these are provided in section 2.

In summary, one can reconstruct the outlines of a linguistically extremely diverse landscape, but available documentation of the non-Quechuan languages leaves much to be wanted and is frequently restricted to placenames and personal names.

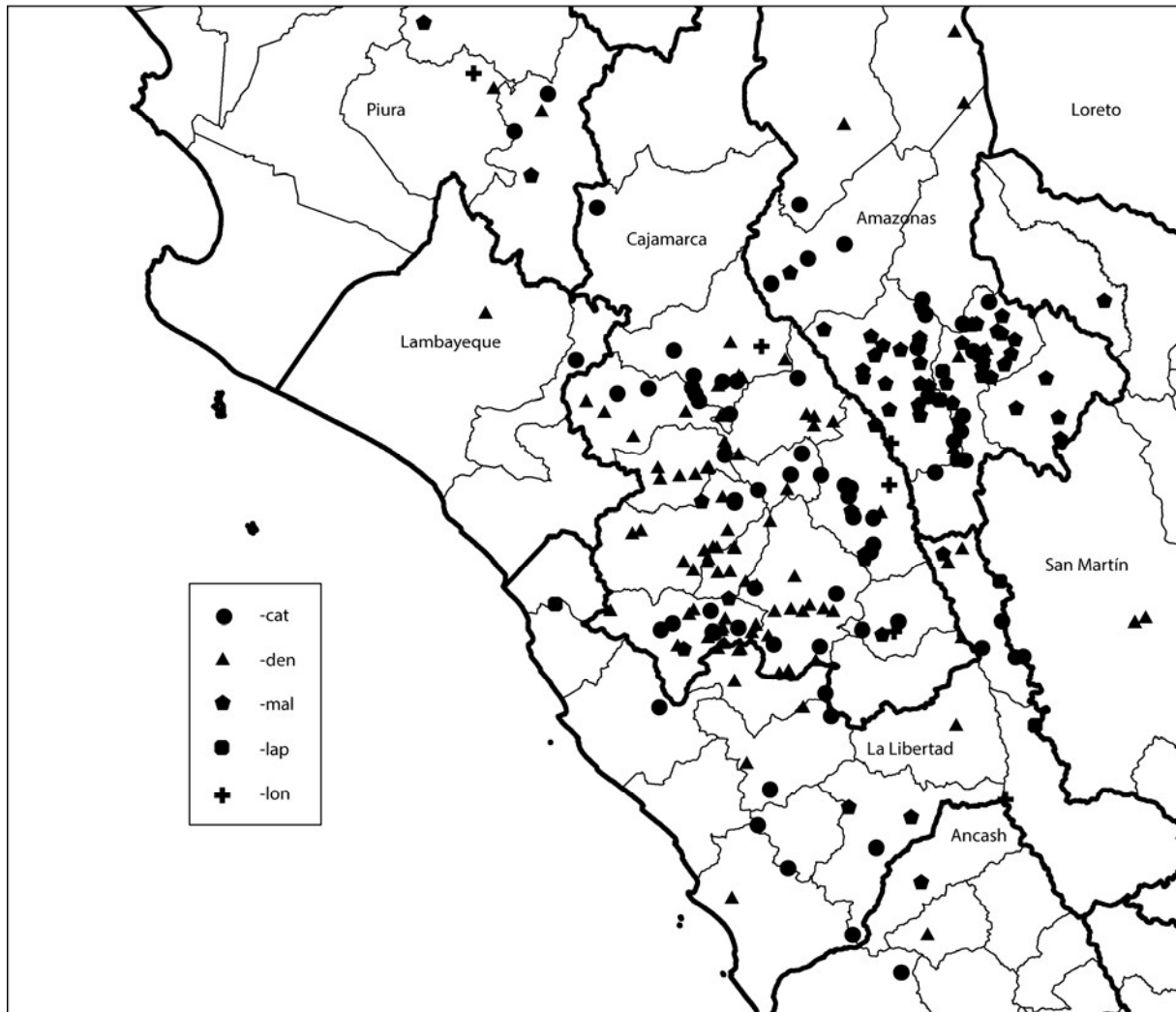


Fig. 1: toponymic areas of the highlands and eastern slopes of northern Peru, based on modern data from the US National Geospatial-Intelligence Agency (<http://geonames.nga.mil/gns/html/>). All variants of endings mentioned in the text have been taken into account for plotting, and some tokens that clearly have a non-indigenous origin were removed post hoc.

Given the absence of crucial pieces of data, especially the linguistic identity and affiliations of the languages known only through toponymy and/or personal names remain obscure and poorly defined. Nevertheless, some suggestive similarities have been pointed out. As far as the highlands and eastern slopes are concerned, one particularly notable observation is that the ending *-cat* is strongly associated with bodies of water throughout its range (Torero 1989: 236). In Chachapoyas, the same semantic link obtains (Valqui Culqui 2004). Accordingly, the ending may be compared with the form <quiet> ‘water’ in Copallín, one of the languages of the Jaén area, and also with *kot*, the Cholón word for ‘water’ (Torero 1989: 236-237, Adelaar with Muysken 2004: 405, Valqui Culqui 2004).<sup>5</sup> Torero (2002: 161), consequently, brings into play the possibility of a genealogical relationship between the Cat language and Cholón.<sup>6</sup>

In this article, I employ the epistemological approach of the triangulation of phenomena (see e.g. Kuorikoski and Marchionni 2016) to advance the understanding of the complex

<sup>5</sup> I use angle brackets for data in premodern sources that are not transcribed according to an explicit transcription system. Conversely, I use the usual italics for linguistic data following a unified transcription system based on phonological considerations (also when they are based on a systematic reconstitution of premodern data, as in the case of Cholón).

<sup>6</sup> There are also wider connections, though; Cullí has <coñ> ‘water’ and <quida> ‘sea’ (Martínez Compañón [1782-1790]1985), Uru *qot[a]* ‘lake’ or similar forms (Hannß 2008: 62), and Quechuan and Aymaran both have forms that go back to the shared etymon \*qutşa (Emlen 2017).

linguistic landscape of the northern Peruvian Andes further. In particular, I suggest that different lines of evidence point to the possibility that one or more “Cholonoid” languages – closely related to Cholón and probably also to Hibito, but possibly distinct from either– could once have covered much wider areas of northern Peru than the area in which Hibito-Cholón is historically known to have been spoken.<sup>7</sup> More specifically, the evidence suggests that the Chacha language as well as the language(s) responsible for the formation of the Den and Cat areas are implicated. Triangulated evidence is of three types: (i) toponymic evidence, building and expanding on pioneering observations by Alfredo Torero, Willem F.H. Adelaar, and Marcelo Jolkesky, (ii) two possible cases of lexical substrate in Chachapoyas Quechua, building on lexicographic and comparative work by Gérald Taylor (1979), and (iii) typological observations on root structure and phonotactics, building on my own work on Central Andean areal typology (Urban 2018, 2019a, b).<sup>8</sup> Each of the types of data comes with unique challenges, strengths, and weaknesses. Nevertheless, together they are consistent in pointing to an affiliation for Chacha, Den, and Cat with Cholón and possibly Hibito.

## 2. An overview of Hibito-Cholón and its speakers

Given the central role of the Hibito and Cholón languages and their speakers in the context of the present article, I provide a somewhat more extensive (but still necessarily incomplete) overview of them in this section. Hibito and Cholón are actually the names given primarily to two historically known ethnic groups of the eastern slopes of the Andes, and only then also to their languages. One particularly early report on the territory inhabited by the groups comes from the diary of Santo Toribio Mogrovejo, a Spanish cleric who travelled widely through Peru at the end of the 16<sup>th</sup> and the beginning of the 17<sup>th</sup> century. In his travel diary, Mogrovejo ([1592-1605]2006) lists the places he visited, and in some cases also provides information on the climate, people, and local languages which he encountered. For many regions of Peru, Mogrovejo’s reports are among the very earliest. Coming from the Chachapoyas area, Mogrovejo ([1592-1605]2006: 141-147) travelled southward from Leymebamba, visiting inter alia Chuquibamba, Uchumarca, Cajamarquilla (modern Bolívar), and Cundumarca (probably Condormarca). The ecclesiastic district of Cundumarca, he says, comprises some villages in mountains called Puymal. Of these, Mogrovejo only visited the main village, Yaro de Puymal. But he mentions also the following further villages, of a province he calls “Zivito”: San Juan de Ulat, Olat, and Abaoto (or Abaotot). The latter, Mogrovejo says in a side remark, marks the boundary between “Zivito” and “Cholón” areas. From there, he travelled further to the village and province of Quisupay, to which also the villages of Nazo, Teputac, Chamal, Laposia, and Suyanti belong. Because of postcolonial resettlements Cholón and Hibito settlements became increasingly interspersed in late colonial times (Poeppig 1836: 321), which is why Mogrovejo’s early report is particularly valuable. But the Cholón-speaking area extended farther south, too. Cholón people are reported as living at the missions of Monzón, Uchiza, Tocache, and Pachiza in the 19<sup>th</sup> century (Poeppig 1836: 321). Poeppig (1836: 327) says that the “Chunchos” –a derogatory generic Quechuan term for forest-dwellers– could easily communicate with speakers of Cholón because their languages were so similar. This could either mean that indeed there was a relative of Cholón still further east, or else that some

<sup>7</sup> “Cholonoid” rather than “Hibito-Cholonoid” or another designation that makes reference to both Hibito and Cholón is chosen because the genealogical relationship between the two, while likely, is not entirely secure (cf. discussion in this section) and the evidence that allows for the identification of the affiliation overwhelmingly comes from Cholón rather than from Hibito.

<sup>8</sup> See Robbeets (2020) for another triangulating approach in which linguistic, archaeological, and genetic data are brought to bear on one another.

speakers of Cholón itself were present east of the Huallaga. Torero (2002: 160-161) in fact mentions an extension of the Cholón language to the jungle of the Ucayali department, citing a statement from the preface to Navarro's (1903: xi) Panoan-Quechua-Spanish dictionary. This implied region is diverse in altitude, climate, and vegetation, ranging from zones of jagged and barren mountainland to humid and hot tropical lowlands.

The historical Cholón were specializing in trade, in particular of salt and coca leaves. Even though their commercial activities were partially instigated by Spanish missionaries (e.g. Sobreviela et al. [1787]1923: 102-103), it is commonly assumed that they were crucial players in facilitating interchange between the Andean highlands and the Amazonian lowlands also in prehistoric times (Reeve 1993: 112-113, Taylor 1999: 217, Torero 2002: 160, cf. Eriksen 2011: 44), a role that befits the location of their land at the intersection of Andes and western Amazonia. As an example for the bridging function of Cholón traders from colonial times, one can mention that Cholón people from Arancai linked Huari in the highlands of Ancash with the shores of the Huallaga (Sobreviela et al. [1787]1923: 102-103). The intermediate position of the Cholón between Andes and Amazonia is also reflected in the structure of their language. This diagnosis has been facilitated in large part through Alexander-Bakkerus's (2005) reconstitution of a 18<sup>th</sup> century colonial grammar by Pedro de la Mata, by far the most extensive source on Cholón. In Cholón there is, on the one hand, a rich case system and a decimal system of numerals like in the major Andean languages of the Quechuan and Aymaran families. On the other hand, there are aspects of Cholón which strongly depart from structures typically found in these languages and which make the language align more with languages of Amazonia (on matters of areal typology see further section 5; see also Urban 2019a for more extensive discussion). Cholón traits that are not typical for most Andean languages, but that are common in Amazonia, include person marking that is prefixal rather than suffixal, and a system of numeral classifiers. These features are exemplified in (1): the allative case marker *-pi* and the Quechuan loanword *ayča* 'meat' can be seen as representing the "Andean" aspects of Cholón, while the classifier *-ta* for firm and/or stony objects and the personal reference system by means of prefixes (here in the form of an auxiliary construction involving a copula) illustrate linguistic structures more common in Amazonia (Alexander-Bakkerus 2005: 271):<sup>9</sup>

- (1) ip- ta- pi ayča Ø- l'up- he i- ton  
 two CL.firm/stony AL meat 3so- eat- SIM 3ps- be.INCOM  
 'Both are eating meat'

Even though "Andean" and "Amazonian" language types are problematic as predefined categories for several reasons, there clearly are traits in Cholón that make it occupy a somewhat ambiguous typological position. This is also visible in typological studies: in the investigation by Urban et al. (2019), which focusses on the western parts of South America and accordingly does not feature Amazonian languages, Cholón clearly clusters with the sampled Quechuan and Aymaran languages. In the study by van Gijn (2014), which does include a range of Amazonian languages, however, Cholón goes together with Arawakan, one of the major language families of Amazonia.<sup>10</sup>

To what extent the grammar of Hibito aligned with either Andes or Amazonia is largely unknown, for no grammar of the language survives. A recurrent final sequence in numerals in the Hibito wordlist of Tessmann (1930: 458-459) suggests the presence of numeral classifiers as in Cholón (Eloranta 2017).

<sup>9</sup> Abbreviations: CL 'classifier', AL 'allative', O 'object', p 'plural', s 'singular', INCOM 'incompletive'.

<sup>10</sup> Note that Torero (2002: 161) suggests contact with Panoan languages.

Lexically, documentation of both Hibito and Cholón is incomplete. While there is no dedicated dictionary of Cholón, a reasonable amount of lexical information has been culled from de la Mata's grammar by Alexander-Bakkerus (2005). The material reveals a moderately strong Quechuan influx, which possibly extends to borrowing of bound morphology as well. Muysken (2012: 239, table 1; 240) thinks of "close contacts in an early period", and suggests that a Quechuan language acted as a "dominant trading language" for speakers of Cholón. In fact, still in the early 20<sup>th</sup> century there apparently were Cholón people on the Ucayali who spoke a variety of Quechua (Navarro 1903: xi).<sup>11</sup> Documentation of Hibito is scarcer and restricted to two wordlists, one found among the famous wordlists of Martínez Compañón ([1782-1790]1985), the other in Tessmann (1930: 458-459). In spite of the very limited material available for Hibito, the genetic nature of the relation between the two languages is widely accepted, even though Torero (1986: 533) is skeptical and considers attributing the lexical similarities to language contact rather than inheritance from a common ancestor. Indeed, 19<sup>th</sup> century sources report intermarriage between Cholón men and Hibito women specifically, an assertion confirmed by present-day descendants of Cholón speakers (Alexander-Bakkerus 2005: 28). While similarities in basic vocabulary are strong and speak in favour of a genetic relationship (Adelaar with Muysken 2004: 461), the social conditions for strong lexical interference – intermarriage and post-contact resettlement blurring the lines between previously distinguishable ethnolinguistic boundaries– were present, so that a non-genetic relationship cannot be entirely discarded.<sup>12</sup>

In spite of missionary activities, Cholón was still viable by the first half of the 19<sup>th</sup> century. Cholón speakers of that time are reported to have spoken a Spanish saliently influenced by Cholón (Poeppig 1836: 327), so that one may assume that the dominant language of these speakers still was Cholón. The language in fact survived until relatively recently. Rememberers consulted by Alexander-Bakkerus (2005) in the 1990s reported that what they knew of Cholón was acquired from their grandparents who still spoke the language.

### 3. Reconsidering the toponymic evidence

#### 3.1. Introduction

In this section, I pursue one line of evidence that suggests a possible Cholón affiliation for some of the extinct languages posited for prehispanic northern Peru. This is toponymy, the study of placenames, which has a high potential for studies aiming to uncover the prior distribution of languages in a given area (cf. Adelaar 2007 and Solís Fonseca 2009 for general perspectives and Cerrón-Palomino 2015 for methodological problems). This, as the discussion in the introductory section has already mentioned, has been masterfully demonstrated for northern Peru by Torero's (1989) study, which traces the distribution of recurrent endings. Especially when overlapping with others that are logically independent, these can be taken to bespeak the presence of a given language in a given area. Rather than establishing toponymic areas on the basis of recurrent endings, here I will often be concerned with the etymology of toponyms as a whole, including recurrent endings, but also the remaining toponymic material where possible. For matters that have mostly to do both with differences in the toponymic record, but also quantitative and qualitative differences regarding prior work in different

<sup>11</sup> This is actually important, because it underscores that Quechuan influence broadly speaking does not necessarily imply highland contact, as a variety of Quechua was and is also a language of the Peruvian (and Ecuadorian) Amazon (cf. Emlen 2016). Nevertheless, as Torero (2002: 161) suggests, it is likely that Cholón also was in contact with Quechua I languages of the highlands.

<sup>12</sup> Poeppig (1836: 329) in fact states that the languages were completely different, though it is unclear how familiar he was with both languages and how much weight should accordingly be given to this statement.

regions, the discussion will have two parts. Section 3.2. treats the Chachapoyas area, section 3.3. the Den and Cat areas in the highlands of Cajamarca and beyond.

### 3.2 Chachapoyas

For the Chachapoyas area, toponymic research begun by Torero (1989) and Taylor (1990) has recently received new impulses by Valqui Cauqui (2004) and Jolkesky (2016: 241, table 10). The former concentrates on inferring the original meaning of toponymic elements through physical characteristics of the very places they denote. The latter presents an intriguing etymologization of Chachapoyas toponymy through Hibito and Cholón lexical material that is one of the starting points of investigation for the present article. Jolkesky takes the earlier observations that the ending *-cat* extends to Chachapoyas and that it might be identified with Cholón *kot* ‘water’ as a starting point. Indeed, incontrovertible toponymic evidence for the presence of Cholón or a related language in regions quite close to the Chachapoyas area exists. For instance, an eastern tributary of the Huallaga river, approx. 130kms southeast of the Chachapoyas town of Leymebamba, appears by the name *Axuacot* in Martínez Compañón’s ([1782-1790]1985) map of Hibito and Cholón missions. This clearly consists of Cholón *ašwa* ‘fish’ and *kot* ‘water’ (Alexander-Bakkerus 2005) or cognates in a closely related language. Regarding the Chachapoyas area proper, Jolkesky goes further in identifying the toponymic ending *-mal*, which is of particular frequency in Chachapoyas, with Cholón *mol* ‘ground, day’. Especially in the sense of ‘ground’ this is a highly plausible item to figure in placenames. In addition, Jolkesky suggests to identify the Chachapoyas ending *-lap* with the Cholón ablative *-(a)p*.<sup>13</sup> As Jolkesky explains in personal communication, the idea is that toponyms of the structure *x-(a)p*, where *x* is an initial sequence that may correspond to a noun, denote a ‘place where *x* exists’ or the like.

However, there are suggestive Chachapoyas place names not mentioned by Jolkesky (2016) which support the presence of a Cholonoid language in the Chachapoyas area as well. Table 2 presents my own attempts to etymologize further placenames of the Chachapoyas region, taken from Valqui Culqui (2004) unless otherwise noted, through Cholón in particular. All Cholón data are cited in the standardized orthography of Alexander-Bakkerus (2005). Criteria were that the Cholón comparanda ought to be plausible candidates to occur in toponyms in terms of their meaning, and that formal differences between toponyms and Cholón comparanda are within reasonable limits.

---

<sup>13</sup> Some comparisons made by Jolkesky would suggest that the */l/* is etymologically part of the preceding root rather than the toponymic ending.



Chachapoyas toponym	Cholón comparison
-mal	<i>mol</i> ‘ground, day’
Colmal	<i>kol</i> ‘death, hunger’
Pangamal ~ Pangomal (?)	<i>pangala</i> ‘forest turkey’ <sup>a</sup>
Sungmal ~ Sugumal	<i>šuj</i> ‘village’, classifier for places, posts, villages, heaps
Laumal	<i>lew</i> ‘caterpillar’
Panamal	<i>pana</i> ‘road’
Yolmal (López de Velasco [1572]1971: 239)	<i>yel</i> ‘salt’
-cat	<i>kot</i> ‘water’
Shúngote	<i>šuj</i> ‘village’, classifier for places, posts, villages, heaps <sup>c</sup>
Muchagache	<i>muč</i> ‘(hot) pepper’ <sup>d</sup>
-ox	
Llondox	<i>ʎom</i> ‘crevice’

Table 1: Chachapoyas placenames with Cholón comparisons

<sup>a</sup> The connection between *pangala* and the toponym Pangamal may be spurious, as the additional syllable of the lexical item is not present in the toponym. Furthermore, this particular item is phonologically anomalous for Cholón, an observation that suggests the possibility of a loan etymology (Alexander-Bakkerus 2005: 109).

<sup>c</sup> It is interesting to note that a place called Huasingate exists in the Chachapoyas area (Valqui Culqui 2004). This apparently is made up of Quechua *wasi* ‘house’ and a variant of the -cat ending; thus, the toponym is closely related semantically to Shúngote if the identification of the initial syllable with Cholón *šuj* is correct.

<sup>d</sup> This appears to be a distinct item from *uč*, which has the same meaning, but is a loan from Quechua. That said, there is the remote possibility of a Quechua etymology for the initial part of the toponym, namely through the verb *mucha-* ‘to kiss’; this item has been borrowed into Cholón as *mučañ* ‘honour, prayer’, surely under missionary influence.

Even though they do not feature traditionally recognized recurrent toponymic endings of the Chachapoyas area, more toponyms of the area can be, if only partially, explained through Cholón. Thus, Opipuy (López de Velasco [1572]1971: 239) can be related to Cholón *pey* ‘earth’, a word that would plausibly form the head of names for landmarks (indeed, we will see that relevant tokens also occur outside the Chachapoyas area). The name of the river Shocol, on the other hand, might be partially explained through Cholón *šo(h)* ‘to pour’ (note that the actual word for ‘river’ in Cholón is *šokot*, literally ‘pouring water’). A particularly interesting case, however, is that of a small town called Limabamba, situated in plain Chachapoyas territory, approximately 50 kilometres southeast of the town of Chachapoyas proper. The etymology of *-bamba* is clear: Quechua *pampa* ‘plain’. The initial sequence may have a Quechua etymology as well: located at the site of a pre-Columbian oracle, after all, the Peruvian capital is related to the Quechua root *rima-* ‘to speak’, and just like the Quechua variety from which the name of Lima derives, Limabamba could reflect the root with /r/ changed to /l/. This may be due to Aymaran influence, as suggested by Cerrón-Palomino (2000) – indeed, a small number of Chachapoyas personal names appear to be Aymaran (Rivarola 2004). Yet, there is a plausible alternative, which is especially attractive because Taylor (1979: 22) mentions that /l/, a few borrowings from other Quechua varieties and cases of spontaneous depalatalization aside, occurs mainly in local words in Chachapoyas Quechua. It is also attractive when considering the geography of Limabamba: as revealed by a satellite map, Limabamba is located on a small area of relatively flat land, befitting the presence of *-bamba*. To its southwest and northeast, this plain is surrounded by mountain chains in the shape of an enclosing horseshoe. Accordingly, Cholón *limañ* ‘mountains, highland’ is a highly

plausible initial element of the placename Limabamba, which would then have a hybrid origin in that its constituents originate from two different languages.<sup>14</sup>

Chachapoyas toponymy, thus, is in some cases explainable through Cholón. However, if the toponymic evidence is taken as an indication of the presence of a Cholonoid language in Chachapoyas, it would be unexpected if the relevant toponymic endings characterized the Chachapoyas area only, but did not occur in the zone in which Cholón is in fact known to have been spoken as well. One would rather expect a continuity of toponymic areas. Indeed, salient Chachapoyan toponymic endings like *-cat* and *-mal* do not cluster sharply in the Chachapoyas region, but extend southward to areas traditionally inhabited by Cholón-speaking people. As far as *-cat* is concerned, the name of the Axuacot river in Martínez Compañón ([1782-1790]1985) is an example; the extension of *-cat* into historical Cholón territory is shown by figure 1. But also *-puy*, which we have just seen in the Chachapoyas toponym *Opipuy*, is an ending which is also attested on the eastern slopes of the Andes much further to the south in what must have been the western margins of the Cholón-speaking area. Note e.g. *Culpuy*, to the east of Cajabamba.

Toponymic continuity is also revealed clearly by historical evidence. As Torero (1989: 238) has observed, Mogrovejo's mentioning of a settlement called Chamal in the Cholón area clearly indicates that the *-mal* toponymic area extends into territory originally inhabited by speakers of Hibito and Cholón, too. The early date of attestation, before the exploitation of the region by the Spaniards at a time when it was a dangerous border zone of Spanish control, virtually excludes post-conquest factors as responsible. Furthermore, an addition to de la Mata's Cholón grammar offers a list of Cholón and Hibito villages, apparently in some cases with Spanish translations. Cholón villages that are mentioned are Apizoncho, Xuñante, Utchinaman, Chalamuy, Chillancuy, Xenquiman, Jallipñatch, Itziuat, Zalcot, and Jopeytè. Xenquiman was first written as Xenquimal, but the <l> then crossed out in favour of <n>. There is thus at least one, and possibly two tokens of *-mal* in an area of the eastern slopes that is known to have been inhabited by speakers of Cholón historically. The translation the addition provides for Jopeytè –‘land like blood’– matches Cholón (*ho* ‘blood’, *pey* ‘land’).<sup>15</sup> This is, at the same time, important evidence for the occurrence of *pey*, attested with the meaning ‘earth’ in the Cholón grammar itself, in a toponym that is not only located in the right area, but that, being identified as a Cholón settlement, is incontrovertibly associated with the Cholón language.

### 3.3. Den and Cat

We have seen that the possible link between placenames in *-cat*, strongly associated with bodies of water, and Cholón *kot* ‘water’ has played a crucial role in suggesting a possible Cholonoid affiliation for the Chacha language of the Chachapoyas region. Also for the Cat area more generally, the connection with Cholón *kot* ‘water’ has been brought up (Torero 1989: 236-237; Adelaar with Muysken 2004: 405). However, similar words for ‘water’ or bodies of water occur widely in the Central Andes, so that Cholón is not necessarily implied specifically. Even if it were, the affiliation with Cholón would not be ultimately secure as a similarity in an isolated form might well also be attributable to chance. Yet, there is more specific evidence that does suggest a connection with Cholón specifically: Torero (1989: 236-

<sup>14</sup> Unless coming from Spanish *limón* ‘lemon’, the element Lima- might recur in the name of a hacienda called Limón, which is located between Bagua Grande and Chachapoyas on Martínez Compañón's ([1782-1790]1985) map of the province of Chachapoyas, as well as in that of another hacienda called Limón, this one near Balsas on the Marañón (Osgood 1914: 155). However, I hesitate to accept these as evidence unless the Spanish-based etymology could be ruled out.

<sup>15</sup> The same is true of Zalcot ‘black river’ (cf. *ʔal* ‘black’, *kot* ‘water’).

237) is able to compare the entire placename Salcot, which occurs repeatedly in Cajamarca, with the name of the Cholón village Zalcot, translated as ‘black water’ in the handwritten addendum to de la Mata’s Cholón grammar (cf. *(ʔi)ʔal* ‘black’ in de la Mata’s grammar, Alexander-Bakkerus 2005). Adelaar (2012: 580) adds another crucial observation, namely that the placename Llacanora of eastern Cajamarca may be explained partially through Cholón *ʔaka* ‘red’. In personal communication, Willem Adelaar in addition points out that Llagadén, another Cajamarcan placename that form part of the Den toponymic area, may contain the same element.<sup>16</sup> Satellite maps actually reveal that both Llacanora and Llagadén are located in close proximity to areas of reddish-ocher land. This is, in addition to the cases pointed out by Adelaar, also true of Llacamate on the western edge of the highlands of La Libertad. Llacamate is located near the wildlife reserve of Calipuy that comprises lands of a similar ocher color (the name Calipuy itself is quite possibly related to Cholón *pey* ‘earth’ as are other placenames in -puy).<sup>17</sup>

Adelaar’s observations are also crucial because they extend the possible Cholón connection to the Den area, too. Accordingly, one can attempt to compare the relevant placenames listed by Torero (1989: 254-257) and to provide etymologies through Cholón in the same manner as done in 3.1. for Chachapoyas. The result is in table 2.

Place	Cholón comparison
Molladén	<i>mol</i> ‘day, ground’
Musadén, Mushadín	<i>mušak</i> ‘sun’
Shuendén	<i>šuy</i> ‘village’, classifier for places, posts, villages, heaps’
Yapodén	<i>yap</i> ‘boar’

Table 2: Additional placenames from the Den toponymic area of Cajamarca and Cholón comparisons.

Adelaar’s suggestion, together with the additional comparisons in table 2, add strength to the more general idea of a prior extension of a Cholonoid languages far into the highlands of northern Peru, Cajamarca in particular. One can note that syllables in relevant Den toponyms are frequently more open, with an additional vowel breaking up consonant clusters that would arise if form like those from Cholón were reflected directly. Etymologizing Cat toponyms through Cholón is, somewhat surprisingly, less productive. Possible comparisons are in table 3.

<sup>16</sup> Torero (1989: 231) states to know of even three instances of places named Llagadén and two instances of Llacadén, for which the same etymology could obviously be proposed.

<sup>17</sup> A 16<sup>th</sup> century report (Anonymous [~1560]1865: 33) on the religious customs of Huamachuco and surroundings, an area where Culli was once spoken, mentions a village named Llaga. In that village, an idol of the same name as well as a collection of deer antlers for apparent magico-religious purposes were found. This conjunction of observations motivated Silva Sansisteban (1985) to suggest that *llaga* is the Culli word for ‘deer’. Even though this need not be the case, the existence of a village of that name in a Culli-speaking area is potentially relevant.

Place	Cholón comparison
Salcot	<i>tʰal</i> ‘black’
Malcat, Melcat, Molecote	<i>mol</i> ‘day, ground’
Puycate	<i>pey</i> ‘earth’ <sup>a</sup>

Table 3: Placenames from the Den toponymic area of Cajamarca and Cholón comparisons.

<sup>a</sup>note Quechua *puyu* ‘cloud’ as an alternative source for the initial syllable in this item

If the connection between Cholón and the Den and Cat areas is accepted (if only as a working hypothesis), one must again ask whether there is any continuity with the toponymy of the known Cholón-speaking area. This would be expected under the interpretation of Den and Cat as Cholonoid languages. Torero (1989: 233) actually has already provided a crucial piece of the answer: from the southern frontier of the Den area as defined by him, there is a line of relevant toponyms that leads to the Marañón valley and also penetrates the high jungle. Relevant placenames are Olmadén ~ Olmadón, Chuquitén, Shuendén, and finally that of the important archaeological site of Pajatén in plain Hibito-Cholón speaking territory which is even “supposed to be an ancient centre of the Cholón or the Hibito” (Alexander-Bakkerus 2005: 27). Torero also notes a number of placenames on the right shores of the Marañón in -én that could be related. And, as we have seen already in section 3.1., placenames in -cat straddle the modern boundary between the departments of La Libertad and San Martín, a region in which historically Cholón was the dominant local language.

### 3.4. Evaluation

Confronted with the puzzle of the relation between Den, Cat, and Culli, Torero (1989: 235) considered that their initial sequences are indicative of very different languages, saying that what coincidences exist (he mentions Aya- in Ayadén and Ayacate), lack diagnostic value. The evidence presented in the preceding section at least weakens that statement. Suggestive similarities are no longer restricted to the ending -cat and Cholón *kot* ‘water’, but extend beyond. Alongside placenames like Limabamba, Llagadén, and Llacanora, three toponymic endings link the historically known Cholón-speaking area with Chachapoyas and large parts of the northern Peruvian highlands. And, as summarized in table 4, all have possible etymologies through Cholón lexical items with meanings that would make their occurrence as recurrent parts of toponyms plausible.

Recurrent toponymic ending	Cholón comparison
-cat	<i>kot</i> ‘water’
-mal	<i>mol</i> ‘ground’
-puy	<i>pey</i> ‘earth, land’

Table 4: Summary of recurrent endings linking the Den and Cat, Chachapoyas, and historically Cholón-speaking areas.

Also, *pace* Torero (1989: 235), apparently Den and Cat do share parts of their lexical stock, as examples like Malcat ~ Melcot ~ Molecote : Molladén or Shúngote: Shuendén show. Importantly, precisely this shared material can be related to Cholón.

At least as an alternative possibility, hence, it is worthwhile to consider Den and Cat not as the signatures of categorically distinct languages, but as somewhat discontinuously distributed toponymic areas that go back to Cholonoid sources that were closely akin to one

another. In addition, the discussion has shown that a considerable number of toponymic endings previously used to identify distinct linguistic areas actually show somewhat overlapping distributions (and, in fact, Torero 1989 had already noted considerable areas of overlap). The distinctiveness of Chachapoyas toponymy is considerable, but also here, there is evidence for continuity of toponymic areas.<sup>18</sup> In the northern Peruvian Andes, the ending in -cat could be considered as defining a continuous Cholonoid toponymic area that includes areas where Cholón itself was dominant in historical times.<sup>19</sup>

#### 4. Substrate lexis

A second pillar of evidence that suggests a prior extension of a Cholonoid language to Chachapoyas in particular comes from a small stratum of Chachapoyas Quechua lexical items which lack good Quechua etymologies. Taylor (1979) has provided the basic necessary steps in investigating this stratum. Not only does he, for each Chachapoyas Quechua item in his dictionary, mention if cognates exist outside Chachapoyas Quechua, but he also provides short appendices listing names for local flora and fauna. Given what is known on the semantic areas in which lexis is likely to be retained from a substrate language, these are good candidates to have outlasted the language shift to Quechua.

One Chachapoyas Quechua lexical item that could stem from a Cholonoid language is *shalla*. Tentatively glossed by Taylor (1979) as ‘basketry’ (“cestería”), possible Quechua parallels are semantically relatively distant. Taylor (1979) himself points out <salla> ‘mat’ in the early Quechua dictionary by Santo Tomás (1560). Nicholas Emlen (p.c.) brings into play Ancash Quechua *shalla*, too, which means ‘thicket’ or ‘weed’, but is perhaps significantly only attested in one of the easternmost provinces of Ancash, namely Antonio Raimondi (Parker and Chávez 1976). There may be a semantic bridge between the meaning attested in Chachapoyas and that given by Santo Tomás (1560) in that baskets are made from straw, just as mats are. Metonymy of a similar type may partially link the Ancash meaning provided by Parker and Chávez (1976).<sup>20</sup>

An alternative etymology for Chachapoyas Quechua *shalla*, however, is through Cholón *šala* ‘basket’, which is much closer semantically than the comparable items within Quechua.<sup>21</sup> The only difference that requires some discussion concerns the place of articulation of the lateral, which is alveolar in Cholón as documented in de la Mata’s colonial grammar, but palatal in Chachapoyas Quechua. In the first place, it is relevant that the alveolar lateral has a somewhat peculiar status in Chachapoyas Quechua according to the description by Taylor (1979: 22). It results from sporadic depalatalization of the palatal lateral, which apparently occurs across varying numbers of lexical items depending on the region. The description provided by Taylor seems compatible with a process of lexical diffusion, i.e. a sound change of depalatalization in progress. Otherwise, as we have already seen in the context of the

<sup>18</sup> Comparable cases of toponymic distributions are known from other parts of the world. The toponymic ending -ingen, for instance, is strongly associated with the High-German speaking areas of southwest Germany and Switzerland (in Bavaria, the variant -ing dominates). However, it is also found in much smaller and less dense clusters in Low German areas (cf. Dahinden 2014: 377, fig. 7). Whatever sociohistorical processes the distribution of -ingen ultimately reflects, it yields discontinuous toponymic areas within the Continental West Germanic dialect continuum.

<sup>19</sup> Befitting its wide distribution, Adelaar with Muysken (2004: 405) call Cat a “toponymic default area” for northern Peru.

<sup>20</sup> Somewhat less clear is how the Junin-Huanca forms *shalla* ‘rubble, rubble field’ and *shala* ‘rocky’ (Cerrón-Palomino 1976) might be related semantically.

<sup>21</sup> Note also that Tsafiki, a Barbacoan language of Ecuador, has *chalá* ‘(kind of) basket, thicket’ (Moore 1966). The relevant basket is known as *chalo* in the Spanish of the region.

toponym Limabamba, the main source of /l/ in Chachapoyas Quechua are pre-Quechua elements (toponyms and terms for flora and fauna). This, as I argue, is precisely the class of items to which *shalla* could belong, too. The depalatalization which is occurring in Chachapoyas Quechua provides in fact another link with Cholón. In the manuscript of de la Mata's Cholón grammar, letters representing the palatal affricate have frequently been crossed out and replaced by alveolars (Alexander-Bakkerus 2005: 88). In some items, this also pertains to laterals. In word-final position, alveolar and palatal points of articulation even appear to have varied freely (Alexander-Bakkerus 2005: 89). Responsible for the emendations that can be observed in the manuscript may be individual variation, dialect differences (many of the corrections seem to have been made not by de la Mata, but by the copyist Gerónimo Clota, who was a missionary in San Buenaventura del Valle, cf. Alexander-Bakkerus 2005: 41, 43) or sound change in progress (Alexander-Bakkerus 2005: 89). In sum, the place of articulation of the laterals apparently varies both in Chachapoyas Quechua and in Cholón diatopically, diachronically, or both. Since in addition at stake here is not Cholón itself but possibly a closely related language, this variation appears to do little damage to the etymology of *shalla* through Cholón.<sup>22</sup> However, a more complex scenario is not excluded. Given the contact history Cholón must have had with Quechua, it is not entirely impossible that Cholón *šala* 'basket' is ultimately from an undefined variety of Quechuan, and was then reintroduced as a borrowing into Chachapoyas Quechua with the specific meaning it had assumed in Cholón.

Another lexical link to Cholón is the Chachapoyas Quechua word *musha*. Some indigenous people of the Chachapoyas region have unusually light skin and blonde or reddish hair, an observation probably already made by the 16<sup>th</sup> century Spanish chronicler Cieza de León (cf. Schjellerup 2005: 61). In Chachapoyas Quechua as well as the Spanish of the Chachapoyas province, such individuals are known as *musha* (Taylor 1979, Malengreau [2008]2015).<sup>23</sup> *Musha* has no plausible cognates in other Quechua varieties, as Taylor (1979) already diagnosed. I believe that a Cholón etymology is highly plausible. Specifically, I identify *musha* with Cholón *mušak* 'sun' (or its cognate in a Cholonoid language). This may seem far-fetched at first glance, but it is not, for there is a common lexico-semantic association between the sun and unusually light-skinned people, including specifically albinos, in indigenous languages of South America of which the one under scrutiny here would just be another instance.<sup>24</sup>

The geographically closest parallel comes from the Mochica language of Peru's north coast. Here, actually, the lexico-semantic association is sensitive to gender and involves both sun and moon: according to Brüning (2004), a male albino is known as <rremik>, while a female albinotic person is referred to by a word transcribed by Brüning as <šang>, <šan'>, <g̃yang>, or <g̃yan'g>. At the same time, the former is the Mochica word for the 'moon', the latter for the 'sun'.<sup>25</sup>

<sup>22</sup> Pertinent fine-grained data on Chachapoyas Quechua and, of course, on different varieties of Cholón are lacking, but it is even a speculative possibility that the two partook in an "depalatalization area" in the last few centuries that would perhaps also include the Quechua of Ancash which registers a fronting of /tʃ/ to [ts]. Torero (2002: 162) explicitly links this process to the variation apparently found in Cholón.

<sup>23</sup> According to Malengreau ([2008]2015), *musha* can also be used to refer to a white person of European descent, while *gringuito* 'little gringo' can be used as an alternative to *musha* for a light-skinned indigenous person.

<sup>24</sup> (Oculocutaneous) albinism of course has to be distinguished from the presence of merely light-skinned and blond phenotypes (otherwise, the Finns would have to be considered albinos). However, since in a South American indigenous context both are highly unusual and phenomenologically similar, one can expect that they may be described in similar terms (this in fact seems to be the assumption of Jeambrun and Sergent 1991, who take traveller's notes on light-skinned people as indications of albinism).

<sup>25</sup> While the word for 'sun' is attested as such in many other sources on Mochica, there is a different word for 'moon' in many other sources. Perhaps importantly, <rremic> is recorded with the meaning 'full moon'

Among the Kuna people of Panama, there is an unusually high incidence of albinism (cf. e.g. Carrasco 2009 for a genetic perspective). For their language, which belongs to the Chibchan family, Holmer (1952) gives *ipekwa* as one of the names for an albino. Holmer also mentions the adverb *ipekwar* ‘like the sun’, and gives *Ipelele* as the name of the sun when thought of as a person. Nordenskiöld with Pérez Kantule (1938: 420-421) states that *ibe* ‘sun’ also applies to albinos directly.<sup>26</sup>

In sum, there are two Chachapoyas Quechua lexical items with unclear or nonexistent Quechuan etymologies that can be plausibly linked with Cholón.<sup>27</sup> The question, however, is if these must necessarily be considered as substrate lexis which survives from the original Chacha language or if borrowing without substrate interference provides a sufficient scenario, in particular given the geographical proximity between Chachapoyas Quechua and Cholón. While such a scenario could account for the presence of *shalla* ‘basket’ in Chachapoyas Quechua, it seems significantly less plausible for *musha*, since this word describes a phenotype that appears very specific to the Chachapoyas region. There would be little reason why a speech community should borrow a word for a highly salient phenotypic phenomenon that occurs among themselves from a contact language such as Cholón.

## 5. Typological considerations

In the preceding sections, I have presented evidence from toponymic distributions and Chachapoyas Quechua vocabulary items that raise the possibility that Cholón or a closely related “Cholonoid” language was once present in large parts of the highlands and eastern slopes of northern Peru. Regarding the toponymic evidence in particular, I have also discussed whether the relevant endings extend to the area where Cholón is in fact known to have been spoken in historical times as a kind of “sanity check” of plausibility. In this section, I pursue such confirmatory lines of reasoning further, but in a somewhat different and logically independent direction. Concretely, if the evidence hitherto presented is genuine and a Cholonoid language indeed once was present in the pertinent areas, then the remains that these Cholonoid languages left in the form of placenames, personal names, substrate lexis, etc. should in terms of phonotactic and syllabic structure, stress patterns etc. be congruent to a large extent to those structures found in Cholón, or at least not be blatantly incompatible with them.

A broad and general perspective on the relevant phenomena in northern Peru is provided in Urban (2019a, b). The general idea in particular of Urban (2019a) is that the dominance in terms of number of individual varieties, number of speakers, and geographic spread of the

---

specifically in the colonial Mochica grammar by Fernando de la Carrera and, as <rem>, with the meaning ‘bright moon’ in the material collected by Ernst Middendorf (Salas García 2002).

<sup>26</sup> Various pieces of evidence from Holmer’s (1952) dictionary point to *ipe* as the original word for ‘sun’ in Kuna, too. *Nakipe*, the current word for the ‘sun’ according to this source, features an additional honorific element *naka*. Another Kuna word for an albino is *sippu* according to Holmer (1952); this word also means ‘white’.

<sup>27</sup> A significantly more speculative commonality could lie in a (fossilized) bound element -um. Valqui Culqui (2004) mentions *Llep* as the name of a local plant in the Chachapoyas area, and suggests that the plant name may be present in the personal name *Llepeum*. The idea that -um is an isolatable element suggested by this observation is reinforced when considering other names ending in -um such as *Milaquium* and *Muchaum* (cf. for the latter also *Muchagache*). Now, -um is a frequent final sequence in *Hibito* and *Cholón* nominals (*Cholón alum* ‘another, other(s)’, *hulum* ‘fatness, thickness’, *kešum* ‘nose’, *lum* ‘top’, *pulum* ‘thunder’, cf. also the *Hibito* forms <puxam> ‘sky’, <coctom> ‘wind’, <chucchum> ‘flower’ in Martínez Compañón [1782-1790]1985). The existence of a fossilized morpheme -um would also support Jolkesky’s (2016) comparison of the Chachapoyas placename *Ulap* with *Cholón owlum* ‘snake’. A final piece of evidence that would suggest that -um is a (historically) separate element comes from the possible comparison of *Hibito* <chucchum> ‘flower’ with *Culli* <chuchú>, which latter lacks the final sequence -um that is putatively specific to *Hibito-Cholón*.

Quechuan family in particular, together with the remarkable contact history this lineage betrays with Aymaran which is discussed prominently in the literature, has led to an excessive weight assigned to the features of that family in theorizing the areal typology of the Central Andes. When the available evidence of the languages of northern Peru is considered, shared structures that in fact differ from typical Quechuan (and Aymaran) patterns become visible. As has first been observed by Torero (2002: 212), the languages of the northern Peruvian Andes documented by Martínez Compañón ([1782-1790]1985), including Hibito and Cholón, show a higher incidence of monosyllabic roots than Quechuan, which has a very strong preference for disyllabic roots. This observation can be extended seamlessly to the personal names of Chachapoyas (Taylor 1990: 124, Valqui Culqui 2004) and also the placenames of the Den and especially the Cat area, in which monosyllabic and disyllabic roots coexist. Furthermore, the northern languages, again including Hibito and Cholón, apparently had less phonotactic restrictions on plosives in word-final position than Quechuan languages, which prohibit /t/ from that position and in which final instances of /p/ are likely due to fossilized suffixation or borrowing from a substrate language (Willem Adelaar p.c.). Again, these regularities extend to Chachapoyas personal names as well as Den and Cat toponyms (cf. e.g. Septén, Cut-Cate). In some basic properties of root structure, thus, the evidence is consistent with a Cholonoid presence in the Den and Cat toponymic areas as well as in Chachapoyas.

However, there are also some differences, of which I would like to discuss two in particular. The first concerns stress. Chachapoyas Quechua has stress on the initial syllable, a highly unusual trait within the Quechuan language family. It is usually assumed that the initial stress is the underlying reason for the characteristic reduction of unstressed vowels in Chachapoyas Quechua. The Chachapoyas Quechua stress shift may be a substrate feature, even though this cannot be demonstrated following conservative lines of reasoning in establishing substrate influence (Thomason 2009) simply because the Chacha stress pattern is not known. If one nevertheless assumes a role of substrate interference in the stress shift, a problematic situation arises since de la Mata describes Cholón as placing stress consistently on the final syllable (Alexander-Bakkerus 2005: 79; the observation is confirmed independently by Poeppig 1836: 327). Tessmann's Cholón data show that pattern, too. So while final stress in Cholón chimes well with the fact that Den toponyms very frequently are stressed on the ending (i.e. the final syllable), explaining the initial stress shift in Chachapoyas Quechua through substratum influence from a Cholonoid language would seem a difficult task.<sup>28</sup> Tessmann's (1930: 458-459) Hibito data, on the other hand, show an inconsistent pattern regarding stress. Not all items are marked for stress; some of those that are show stress on the last, others on the first syllable (e.g. <montsá> ~ <mantas> 'eye', <sótša> 'head').

Another observation worthy of discussion is the presence of letters <b>, <d>, and <g> in initial position of Chachapoyas personal names. These suggest that voicing in stops was phonologically contrastive in Chacha (Taylor 1990). The Cholón grammarian de la Mata, in contrast, clearly states that "[i]n this language, the letters B, D, F and R, [...] are not pronounced" (Alexander-Bakkerus 2005: 51fn2). Again, Tessmann's Hibito data show some more evidence of voiced stops and affricates than his Cholón data, but the differences are neither dramatic nor consistent. One noteworthy item is the word for 'tooth', given as <dzuī> with the alternative transcription <tui> in parenthesis. Bearing in mind that letters <b>, <d>, and <g> could also have been used to represent something different from [b], [d], [g], as also happened in an unsystematic rendering of the Barbacoan language Guambiano using the

<sup>28</sup> Another observation that would suggest the possibility of a particularly close link of Den with Cholón is the absence of evidence for [w] in both: in all Den toponyms in Torero (1989) orthographic sequences that would strongly suggest the presence of that sound are absent.



Spanish alphabet (cf. Urban submitted), the inconsistency regarding voicing in stops between Chacha personal names and placenames remain.

In particular as far as Chachapoyas is concerned, then, there is evidence for some typological discrepancies between what traits can be tentatively posited for the original language of the region and Cholón and to a lesser extent also Hibito. On other levels of analysis, namely root structure and phonotactic structure, the data reinforce the ties of the languages that created the Den and Cat areas, as well as Chacha, with the northern Peruvian Andes generally and also Cholón specifically.

## 6. Discussion

The available linguistic evidence for many languages of the highlands is extremely restricted, and leaves the researcher wanting for more. What little linguistic evidence there is, however, is consistent with a scenario in which languages closely related to Cholón were once present in large parts of the highlands of northern Peru, in particular Cajamarca. Needless to say, the type of analysis that is possible on the basis of the available data does not even approach the security that can be attained regarding a genealogical link by the application of the comparative method. Even if the strength of the different lines of evidence were stronger, which is logically possible and might even be achieved by further research (more possible etymologies of Den and Cat toponyms through Cholón, more Chachapoyas lexical items that can be linked to Cholón, and stronger typological affinities), by the nature of the evidence the inference of a Cholonoid extension into the highlands of Chachapoyas and Cajamarca would remain a matter of probabilities only. Neither now nor in the future should the idea of a larger extension of the Hibito-Cholón family be accepted uncritically as fact. Also, the discovery of additional relevant data, such as dedicated documentation of the Chacha language or of Hibito, could necessitate a reconsideration of the posited scenario. Moreover, in any case it is necessary to not only consider the spatial dimension, but also the temporal dimension, a point recently made with reference to the Peruvian North by Andrade Ciudad (2010). The point that the Den, Cat, and Chachapoyas toponymy could indicate the presence of a Cholonoid language at certain times in prehistory does not necessarily mean that at any one point of time such a language was the dominant language of the entire area. It is possible that the toponymic signatures pertain to different temporal strata, as Andrade Ciudad (2010) in fact argues. This possibility requires further consideration in the light of the evidence presented in this article.

But even when focusing for the time being on spatial rather than temporal distributions alone, the evidence presented here makes broader (re)considerations of a multidisciplinary nature possible. Valqui Cauqui (2004) notes the mismatch between the hitherto established extension of the Chacha language as inferred by toponymy and that of the Chachapoyas culture. They overlap reasonably in the north, but the area affiliated with the Chachapoyas culture extends considerably farther southward into the Marañón-Huallaga interfluvium, with sites such as Nunamarca just north of Tayabamba (Church and von Hagen 2008: 905, fig. 45.1). To be sure, modern research does not expect a one to one match between linguistic and cultural distributions anymore as was largely the case until well into the 20<sup>th</sup> century. Since archaeological styles are no longer interpreted as the product of a homogeneous people, and since these in turn need not have been linguistically homogeneous, such a mismatch is in principle unproblematic. In this case, nevertheless, the possibility of a prior extension of a Cholonoid language to Chachapoyas does yield a remarkable congruence between archaeological-cultural and linguistic data.

## References

- Adelaar, Willem F.H. 1988. Search for the Culli language. Continuity and identity in Native America. Essays in honor of Benedikt Hartmann, ed. by Maarten Jansen, Peter van der Loo, and Roswitha Manning, 111-131. Leiden/New York/København/Köln: E.J. Brill.
- Adelaar, Willem F.H. 2007. The importance of toponymy, family names and historical documentation for the study of disappearing and recently extinct languages in the Andean region. Language endangerment and endangered languages. Linguistic and anthropological studies with special emphasis on the languages and cultures of the Andean-Amazonian border area, ed. by W. Leo Wetzels, 325-331. Leiden: CNWS.
- Adelaar, Willem F.H. 2012. Languages of the Middle Andes in areal-typological perspective: emphasis on Quechuan and Aymaran. The indigenous languages of South America: a comprehensive guide, ed. by Lyle Campbell and Verónica Grondova, 575-624. Berlin/Boston: Walter de Gruyter.
- Adelaar, Willem F.H., with the collaboration of Pieter C. Muysken. 2004. The languages of the Andes. Cambridge: Cambridge University Press.
- Alexander-Bakkerus, Astrid. 2005. Eighteenth-century Cholón. Utrecht: LOT.
- Andrade Ciudad, Luis. 2010. Contactos y fronteras de lenguas en la Cajamarca prehispánica. Lenguas y sociedades en el antiguo Perú. Hacia un enfoque interdisciplinario, ed. by Peter Kaulicke, Rodolfo Cerrón-Palomino, Paul Heggarty, and David Beresford-Jones. Special issue of *Boletín de Arqueología* 14.165-180.
- Anonymous [~1560]1865. Relación de la religión y ritos del Perú, hecha por los primeros religiosos Agustinos que allí pasaron para la conversión de los naturales. Colección de documentos ineditos relativos al descubrimiento, conquista y colonización de las posesiones españolas en América y Oceanía, sacados, en su mayor parte, del Real Archivo de Indias, vol 3, ed. by Joaquin F. Pacheco, D. Francisco de Cárdenas, and D. Luis Torres de Mendoza, 5-58. Madrid: Manuel B. de Quirós.
- Brüning, Hans Heinrich. 2004. Mochica Wörterbuch. Diccionario mochica. Mochica – castellano / castellano – mochica. Ed. by José Antonio Salas García. Lima: Universidad de San Martín de Porres, Escuela Profesional de Turismo y Hotelería.
- Carrasco, Androuw, Elaine M. Forbes, Pascale Jeambrun, and Murray H. Brilliant. 2009. A splice site mutation is the cause of the high prevalence of oculocutaneous albinism type 2 in the Kuna population. *Pigment Cell and Melanoma Research* 22/5.645-647.
- Cerrón-Palomino, Rodolfo. 1976. Diccionario quechua. Junín-Huanca. Lima: Ministerio de Educación/Instituto de Estudios Peruanos.
- Cerrón-Palomino, Rodolfo. 2000. Nota etimológica. El topónimo Lima. *Lexis* 24/1.151-162.
- Cerrón-Palomino, Rodolfo. 2004. Lenguas de la costa norte peruana. Estudios en lenguas amerindias. Homenaje a Ken L. Hale, ed. by Zarina Estrada Fernández, Ana V. Fernández Garay, and Albert Álvarez González, 81-105. Hermosillo: Editorial Unison.
- Cerrón-Palomino, Rodolfo. 2015. Toponimia andina. Problemas y métodos. *Lexis* 39/1.183-197.
- Church, Warren B., and Adriana von Hagen. 2008. Chachapoyas. Cultural development at an Andean cloud forest crossroads. Handbook of South American archaeology, ed. by Helaine Silverman and William H. Isbell, 903-926. New York: Springer.
- Dahinden, Tobias. 2014. Towards a spatial analysis of toponym endings. Cartography from pole to pole, ed. by Manfred Buchroithner, Nikolas Prechtel, and Dirk Burghardt, 369-379. Heidelberg/New York/Dordrecht/London: Springer.
- Eloranta, Rita. 2017. Language contact across the Andes. The case of Mochica and Hibito-Cholón. Language contact and change in Mesoamerica and beyond, ed. by Karen Dakin,

- Claudia Parodi, and Natalie Operstein, 319-334. Amsterdam/Philadelphia: John Benjamins.
- Emlen, Nicholas Q. 2016. Multilingualism in the Andes and Amazonia. A view from in-between. *Journal of Latin American and Caribbean Anthropology*. doi:10.1111/jlca.12250.
- Emlen, Nicholas Q. 2017. Perspectives on the Quechua-Aymara contact relationship and the lexicon and phonology of Pre-Proto-Aymara. *International Journal of American Linguistics* 83/2.307-340.
- Eriksen, Love. 2011. Nature and culture in prehistoric Amazonia. Using G.I.S. to reconstruct ancient ethnogenetic processes from archaeology, linguistics, geography, and ethnohistory. Dissertation, Lund University.
- Hannß, Katja. 2008. Uchumataqu. The lost language of the Urus of Bolivia. A grammatical description of the language as documented between 1894 and 1952. Leiden: CNWS.
- Holmer, Nils M. 1952. Ethno-linguistic Cuna dictionary, with indices and references to *A critical and comparative Cuna grammar* (Etnologiska studier 14) and the *Grammatical sketch in Cuna chrestomathy* (Etnologiska Studier 18). Göteborg: Etnografiska Museet.
- Jeambrun, Pascale, and Bernard Sergent. 1991. Les enfants de la lune. L'albinisme chez les Amérindiens. Paris: INSERM-ORSTOM.
- Jolkesky, Marcelo Pinho de Valhery. 2016. Estudo arqueo-ecolinguístico das terras tropicais sul-americanas. Dissertation, Universidade de Brasília.
- Kuorikoski, Jaakko, and Caterina Marchionni. 2016. Evidential diversity and the triangulation of phenomena. *Philosophy of Science* 83/2.227-247.
- López de Velasco, Juan. [1572]1971. Geografía y descripción universal de las Indias. Ed. by Marcos Jiménez de la Espada. Madrid: Ediciones Atlas.
- Malengreau, Jacques. [2008]2015. Parientes, paisanos y ciudadanos en los Andes de Chachapoyas. Identidades, divisiones sociales y solidaridad en la comunidad de San Carlos. Lima: Institut français d'études andines. Online: <http://books.openedition.org/ifea/5495>
- Martínez, Compañón, Baltasar Jaime. [1782-1790]1985. Trujillo del Perú en el siglo XVIII. Vol. 2. Madrid: Ediciones Cultura Hispánica.
- Mogrovejo, Toribio. [1593-1605]2006. Libros de visitas de Santo Toribio Mogrovejo (1593-1605). Ed. by José Antonio Benito. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú.
- Moore, Bruce R. 1966. Diccionario castellano-colorado, colorado-castellano. Quito: Instituto Lingüístico de Verano.
- Muysken, Pieter. 2012. Contacts between indigenous languages in South America. The indigenous languages of South America. A comprehensive guide, ed. by Lyle Campbell and Verónica Grondona, 235-258. Berlin/Boston: Walter de Gruyter.
- Navarro, Manuel. 1903. Vocabulario castellano – quechua – pano, con sus respectivas gramáticas quechua y pana. Lima: Imprenta del Estado.
- Nordenskiöld, Erland, in collaboration with Ruben Pérez Kantule. 1938. An historical and ethnological survey of the Cuna indians. Ed. by Henry Wassen. Göteborg: Göteborgs Museum, Etnografiska Avdelningen.
- Osgood, Wilfred H. 1914. Mammals of an expedition across northern Peru. *Field Museum of Natural History, Zoological Series* 10/12.143-185.
- Parker, Gary J., and Amancio Chávez. 1976. Diccionario quechua. Ancash-Huailas. Lima: Ministerio de Educación/Instituto de Estudios Peruanos.
- Poeppig, Eduard. 1836. Reise in Chile, Peru und auf dem Amazonenstrome, während der Jahre 1827-1832, vol. 2. Leipzig: Friedrich Fleischer/J. C. Hinrichssche Buchhandlung.

- Reeve, Mary-Elizabeth. 1993. Regional interaction in the Western Amazon. The early colonial encounter and the Jesuit years. 1538-1767. *Ethnohistory* 41/1.106-138.
- Rivarola, José Luis. 2004. Antroponimia chachapoya. *La Romania en interacción. Entre historia, contacto y política. Ensayos en homenaje a Klaus Zimmermann*, ed. by Martina Schrader-Kniffki and Laura Morgenthaler García, 295-305. Madrid/Frankfurt am Main: Iberoamericana/Vervuert.
- Robbeets, Martine. 2020. The Transeurasian homeland. Where, what and when?, *The Oxford Guide to the Transeurasian Languages*, ed. by Martine Robbeets and Alexander Savelyev, 772-783 Oxford: Oxford University Press.
- Rostworowski de Diez Canseco, María. 1985. Patronyms with the consonant F in the *Guarangas* of Cajamarca. *Andean ecology and civilization. An interdisciplinary perspective on Andean ecological complementarity*, ed. by Shozo Masuda, Izumi Shimada, and Craig Morris, 401-421. Tokyo: University of Tokyo Press.
- Salas García, José Antonio. 2002. *Diccionario mochica–castellano*. Lima: Universidad de San Martín de Porres, Escuela Profesional de Turismo y Hotelería.
- Salas García, José Antonio. 2010. La lengua Pescadora. *Boletín de la Academia Peruana de la Lengua* 50.83-128.
- Santo Tomás, Domingo de. 1560. *Vocabulario de la lengua general de los indios del Perú, llamada Quichua*. Valladolid: Francisco Fernandez de Cordoua.
- Schejellerup, Inge R. 2005. *Incas y españoles en la conquista de los chachapoya*. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú/Instituto Francés de Estudios Andinos.
- Silva Sansisteban, Fernando. 1985. La lengua culle de Cajamarca y Huamachuco. *Historia de Cajamarca, vol. 2: etnohistoria y lingüística*, ed. by Fernando Silva Sansisteban, Waldemar Espinoza Soriano, and Rogger Ravines (eds.), 365-369. Cajamarca: Instituto Nacional de Cultura.
- Sobreviela, Manuel, Francisco Álvarez de Villanueva, and Vicente Gómez [1787]1923. *Diario del viaje que hicimos a las conversiones de las fronteras de Huánuco partido de la intendencia de Tarma arzobispado de Lima y de Pataz o Cajamarquilla, intendencia y obispado de Trujillo el padre guardián del Colegio de Propaganda Fide de Santa Rosa de Ocopa Fr. Manuel Sobreviela, el padre ex-comisario Fr. Francisco Álvarez de Villanueva [sic!] y el padre secretario de la visita Fr. Vicente Gómez en el año de 1787. Historia de las misiones franciscanas y narración de los progresos de la geografía en el oriente del Perú. Relatos originales y producciones en lenguas indígenas de varios misioneros*, ed. by Bernardino Izaguirre, 77-155. Lima: Talleres Tipográficos de la Penitenciaría.
- Solís Fonseca, Gustavo. 2009. No todo desaparece cuando una lengua muere. *Construyendo Nuestra Interculturalidad* 5/4.1-11.
- Stout, David B. 1948. The Cuna. *Handbook of South American Indians, vol. 4: the circum-Caribbean tribes*, ed. by Julian H. Steward, 257-268. Washington: United States Government Printing Office.
- Taylor, Anne Christine. 1999. The western margins of Amazonia from the early sixteenth to the early nineteenth century. *The Cambridge history of the native peoples of the Americas, vol. 3.2*, ed. by Frank Salomon and Stuart B. Schwartz, 188-256. Cambridge: Cambridge University Press.
- Taylor, Gérald. 1979. *Diccionario normalizado y comparativo quechua. Chachapoyas-Lamas*. Paris: L'Harmattan.
- Taylor, Gérald. 1990. La lengua de los antiguos Chachapuyas. *Temas de lingüística amerindia. Primer Congreso Nacional de Investigaciones Lingüístico-Filológicas*, ed. by

- Rodolfo Cerrón-Palomino and Gustavo Solís Fonseca, 121-139. Lima: Consejo Nacional de Ciencia y Tecnología/ Programa de Educación Bilingüe de Puno.
- Tessmann, Günter. 1930. *Die Indianer Nordost-Perus. Grundlegende Forschungen für eine systematische Kulturkunde*. Hamburg: Friederichsen, De Gruyter & Co. M. B. H.
- Thomason, Sarah G. 2009. How to establish substratum influence. *Issues in Tibeto-Burman Historical Linguistics*, ed. by Yasuhiko Nagano, 319-328. Osaka: National Museum of Ethnology.
- Torero, Alfredo. 1986. Deslindes lingüísticos en la costa norte peruana. *Revista Andina* 4/2.523-548.
- Torero, Alfredo. 1989. Areas toponímicas e idiomas en la sierra norte peruana. Un trabajo de recuperación lingüística. *Revista Andina* 7/1.217-257.
- Torero, Alfredo. 1993. *Lenguas del nororiente peruano. La hoya de Jaén en el siglo xvi*. *Revista Andina* 11/2.447-472.
- Torero, Alfredo. 2002. *Idiomas de los Andes. Lingüística e historia*. Lima: Instituto Francés de Estudios Andinos/Editorial Horizonte.
- Urban, Matthias. 2019a. Is there a Central Andean linguistic area? A view from the perspective of the “minor” languages. *Journal of Language Contact* 12/2.271-304.
- Urban, Matthias. 2019b. *Lost languages of the Peruvian North Coast*. Berlin: Ibero-American Institute/ Gebr. Mann.
- Urban, Matthias. 2018. The lexical legacy of substrate languages. A test case from the Southern Ecuadorian highlands. *Transactions of the Philological Society* 116/3.435-459.
- Urban, Matthias, Hugo Reyes-Centeno, Kate Bellamy, and Matthias Pache. 2019. The areal typology of western Middle and South America. Towards a comprehensive view. *Linguistics* 52/6.1403-1463.
- Valqui Culqui, Jairo. 2004. *Reconstrucción de la lengua chacha mediante un estudio toponímico en el distrito de La Jalca grande (Chachapoyas-Amazonas)*. Licenciante thesis, Universidad Nacional Mayor de San Marcos.
- Van Gijn, Rik. 2014. The Andean foothills and adjacent Amazonian fringe. The native languages of South America. Origins, development, typology, ed. by Loretta O’Connor and Pieter Muysken, 102-125. Cambridge: Cambridge University Press.
- Zevallos Quiñones, Jorge. 1966. Onomástica prehispánica de Chachapoyas. *Lenguaje y Ciencias* 20.27-41.