Contributions to Chibchan historical linguistics

Pache, M.J.

1 Introduction

Chibchan has been labelled as "by all means one of the most intriguing language groups in the Americas" (Adelaar 2007: 11), and for western researchers, this appears to be particularly true in the domain of historical linguistics: Several different authors have contributed to the reconstruction of Proto-Chibchan forms, although with sometimes different results (cf., e.g., Wheeler 1972; Holt 1986; Constenla Umaña 1981, 2012), and many have investigated the possible external classification of the family, a puzzle that has also not been conclusively put together (cf., e.g., Herzog 1886; Uhle 1890; Jijón y Caamaño 1943; Constenla Umaña 2005).

The present thesis aims to contribute to this ongoing discussion. It is divided into two parts:

- First, it will discuss family-internal issues such as the phonology and morpheme-inventory of Proto-Chibchan (section 2).
- Second, it will discuss the external classification of Chibchan (section 3). This second part of the thesis is based on the Proto-Chibchan reconstruction provided in section 2.

Chibchan languages are spoken at the very heart of the Americas, on the isthmus connecting both continents and adjacent regions, also known as 'Intermediate Area' (cf. Haberland 1957: 156, cited in Constenla Umaña 1991: 5). More specifically, Chibchan languages are distributed across a distance of about 1,500 km on an east–west axis, from northwestern Venezuela to northeastern Honduras (see figure 1). Because of this geographic position, the Chibchan family has been granted a particular importance in the investigation of prehistoric intercontinental relationships (Mason 1950: 175; Constenla Umaña 1981: 17).

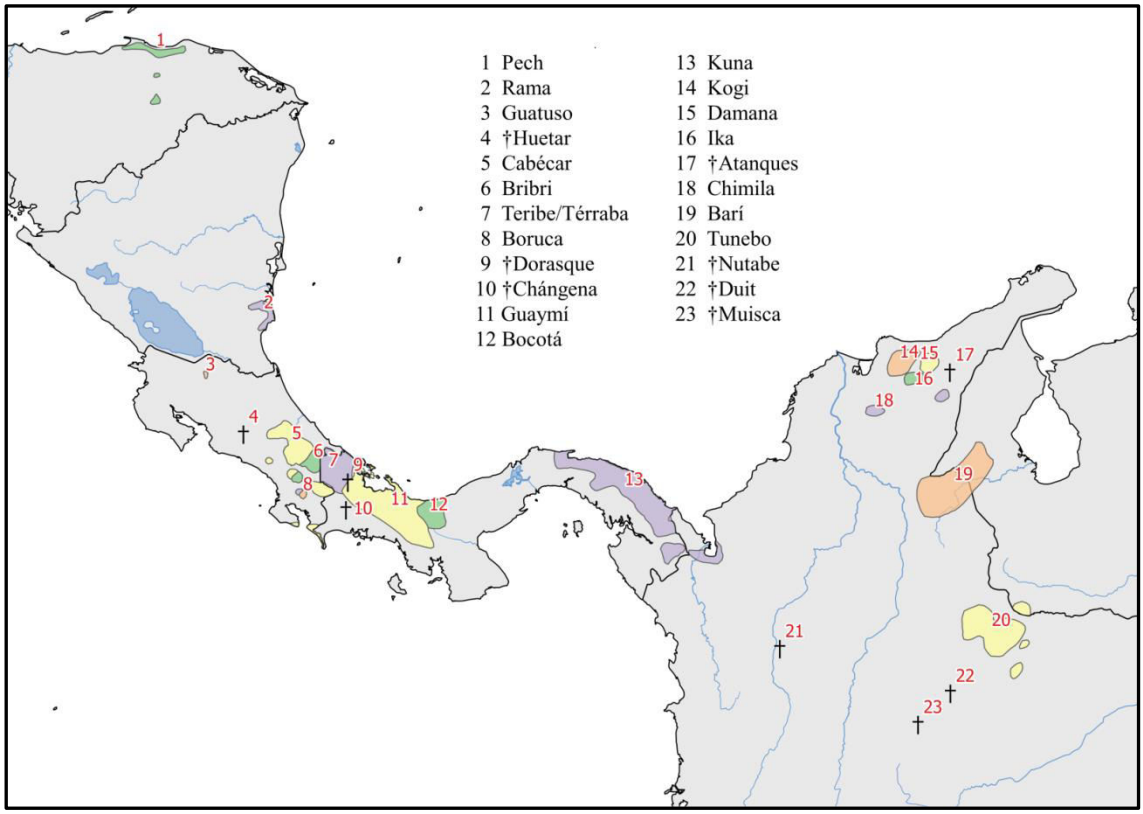


Figure 1: The distribution of Chibchan languages. The map was created by Arjan Mossel, based on data from Constenla Umaña (2012: 394) and Simons and Fennig (2017).

The existence of the Chibchan family was recognized in the late nineteenth century by European researchers (Herzog 1886; Uhle 1890), and since the 1940's (cf. Holmer 1947: 207–8), reconstructive work has been undertaken. The most comprehensive so far are the Proto-Chibchan reconstructions of Constenla Umaña (1981) and of Holt (1986). The authors do not always agree, however: For instance, no Proto-Chibchan mid vowels have been reconstructed by Holt (1986), who postulates the existence of only four Proto-Chibchan vowels, *i, *ɨ, *u, and *a. In contrast, Constenla Umaña (2012) reconstructs ten vowels, *i, *ĩ, *u, *ũ, *e, *ẽ, *o, *õ, *a, *ã, that is, a Proto-Chibchan set which does include mid vowels. It goes without saying that several differences also exist in the domain of the reconstructed lexical forms. Some of these differences might be due to the fact that Proto-Chibchan, as reconstructed by Constenla Umaña (1981) and Holt (1986), is based on data from different Chibchan languages, six in both cases, but not exactly the same ones. This suggests that the phonology and the morpheme inventory of Proto-Chibchan need to be reconsidered on the basis of data from more

Chibchan languages. Meanwhile, many descriptive works on Chibchan have become available since the 1980's, after the time when Constenla Umaña and Holt had prepared their theses (e.g., Chevrier (2017aa) on Bribri; a comprehensive list of the works in question is given below, in subsections 2.1.2 and 2.3). Thus, this is now an opportune moment for (1) an amplification of existing reconstructions, for (2) a reconsideration of the reflexes of Proto-Chibchan in single Chibchan languages forms proposed so far, and (3) for an updated reconstruction of Proto-Chibchan phonology, based on recently published sources. The present thesis proposes a revised, thorough and updated reconstruction of Proto-Chibchan, based on recent publications and on my own fieldwork in the case of Barí, a little-known Chibchan language of the Colombian–Venezuelian borderland region. This is the first time that a reconstruction of Proto-Chibchan is based on data from all twenty-three Chibchan languages that are known at present.

Whereas reconstructive work has been a topic of interest since the 1940's, the external classification of Chibchan has been debated since the late nineteenth century. The proposals discussing the external classification of Chibchan since then may grossly be divided into two groups: (1) those implying connections with northern languages (Mesoamerica, northern Central America) and (2) those implying South American connections. Among the linguists who put forward genealogical connections of Chibchan with Mesoamerican languages such as Uto–Aztecan are Swadesh (1954: 324–5) and Holt (1997/98). Connections with the Central American language groups Misumalpan and Lencan were proposed by Mason (1950), Greenberg (1987), and, above all, Constenla Umaña (2002, 2005, 2012). This putative connection of Chibchan with Lencan and Misumalpan was recently subsumed under the label of a 'Lenmichí Microphylum' (Constenla Umaña 2012: 418) and has possibly been the most influential proposal in the recent past. Yet, Lenmichí has not been unanimously accepted as a demonstrable genealogical unit by all researchers either (cf. Campbell 2012: 134). Instead, genealogical links of Chibchan with languages spoken far into South America have been proposed from early on by several different researchers (e.g., Herzog 1886; Rivet 1924a; Jijón y Caamaño 1943; Loukotka 1968; Greenberg 1987). For instance, a genealogical connection has been proposed to exist between Chibchan and the Páez and Andakí languages of southern Colombia (e.g., Rivet 1924a: 680–4), or between some Chibchan languages of Central America and the Cariban languages of eastern South America (Herzog 1886). Although many different South American languages have been connected to Chibchan languages without any conclusive result, these linguistic proposals seem to concur with cultural parallels and genetic coincidences between Chibchan-speaking groups and South American populations (cf., e.g., Kidder 1940:

444–5; Wang et al. 2007). As a second contribution to Chibchan historical linguistics, the present thesis discusses the external classification of Chibchan.

Before dealing with the reconstruction of Proto-Chibchan and the external classification of this family in the main part of this study, the reader will find some background information in the following subsections. I will briefly introduce Chibchan languages and peoples in subsection 1.1, and some background information on Chibchan historical linguistics in subsection 1.2. The introduction closes with an overview of the structure of this thesis (1.3).

1.1 Chibchan languages and peoples

In the following subsections, I will give a brief overview of Chibchan-speaking groups in terms of languages (1.1.1), culture and archaeological findings (1.1.2) and genetic particularities (1.1.3)

1.1.1 Languages

The Chibchan family is comprised of some twenty-three languages. They are briefly presented in table 1, based on Quesada (2007: 34–5), together with present-day countries in which they are or were spoken, and speaker numbers.

The first description of a Chibchan language was published in the early seventeenth century, Lugo's (1619) Muisca grammar and catechism. It remained the only published source on a Chibchan language for more than two hundred years. Early works on other Chibchan languages and catechisms for clergymen seem to have circulated in manuscript form during colonial times (cf. Lehmann 1920: 236; Rey Fajardo 1971). With a few exceptions however (e.g., Pinart 1882; Villamañán 1978; Lucena Salmoral 1964–69; González de Pérez 1987), these manuscripts were never edited and seem to have been lost or still await to be rediscovered in some archive. At present, published materials dealing with phonology and/or morphosyntax are available for all spoken Chibchan languages (for more information, see subsection 2.3). Notwithstanding, several Chibchan languages are still understudied, such as Pech in Honduras, and some Chibchan languages of northern South America, such as Barí, to mention just a few examples. Overviews of the phonological and morphosyntactic features of South American Chibchan languages are given by Adelaar and Muysken (2004), and different contributors in González de Pérez and Rodríguez de Montes (2000), such as Landaburu (2000) or Mogollón Pérez (2000). Quesada (2007) covers Chibchan languages of both South and Central America.

TABLE 1
OVERVIEW OF CHIBCHAN LANGUAGES

Language	Country	Number of speakers
Atanques, Kankuí, Kakatukua	Colombia	Extinct (?)
Barí, Motilón	Colombia, Venezuela	2,500
Bocotá, Buglere	Costa Rica, Panama	2,500
Boruca, Brunca	Costa Rica	1
Bribri	Costa Rica	6,000
Cabécar	Costa Rica	2,500
Chánguena	Panama	Extinct
Chimila, Ette taara	Colombia	2,000
Damana, Arsario, Sanká, Malayo, Marocasero, Guamaca	Colombia	2,800
Dorasque: Chumulu and Gualaca	Panama	Extinct
Duit	Colombia	Extinct
Guatuso, Maleku	Costa Rica	365
Guaymí, Ngäbe, Ngäbére, Movere, Move	Costa Rica, Panama	150,000
Huetar	Costa Rica	Extinct
Ika, Bintucua, Arhuaco	Colombia	8,000
Kogi, Kogui, Cágaba	Colombia	7,000
Kuna	Colombia, Panama	70,000
Muisca, Chibcha	Colombia	Extinct
Nutabe	Colombia	Extinct
Pech, Pesh, Paya	Honduras	600
Rama	Nicaragua	24
Térraba and Teribe, Naso	Costa Rica, Panama	1,500
Tunebo, Uw Cuwa	Colombia, Venezuela	1,800 (Colombia)

Data from Quesada (2007: 34–5).

The degree of vitality of each of the Chibchan languages varies a lot. Whereas its westernmost members Pech, Rama, and Guatuso are heavily endangered or virtually extinct at present, other members of the family, such as Guaymí and Kuna are still spoken by a fair amount of people (more than 150,000 speakers in the case of Guaymí, according to Quesada Pacheco 2008: 15; see also table 1 above).

In terms of typological profiles, the Chibchan family is likewise diverse. Noticeable differences between Central American and Colombian Chibchan languages are found in the case of syllable-initial *muta-cum-liquida* consonant clusters (such as *kr* or *kɣ*), which are frequent in some Central American Chibchan languages, but barely occur in Chibchan languages of Colombia (see below, subsections 2.3 and 2.4.4.1). A similar pattern of distribution exists for vowel nasality spreading to neighbouring consonants in some Chibchan languages of Central America (e.g., in Bribri and Cabécar (Constenla Umaña 1985a; Chevrier 2017a).

In the domain of morphology there is likewise a high degree of heterogeneity in Chibchan. Free forms, prefixes, suffixes or a mix of both may indicate verbal person in different languages (Pache 2015). Verbal morphology can be very complex in terms of irregularities, participant-encoding, or epistemic marking in certain Colombian Chibchan languages such as Muisca, Ika, or Kogi (cf. Landaburu 1992; Adelaar & Muysken 2004: 89–103; Bergqvist 2016). In contrast, the verbal morphology of Kuna appears to be comparatively straightforward (cf. Smith 2014).

1.1.2 Culture and archaeology

From a cultural point of view, Chibchan-speaking groups are likewise heterogeneous, sometimes displaying features that were typically associated with cultures from the Central Andes or of the tropical lowlands of South America (e.g., by Steward 1948). For instance, one may tentatively identify the following 'Lowland South American' features among the Barí, a Chibchan group of the Colombian-Venezuelian border area (e.g., Beckerman & Lizarralde 2014: 216–30): the use of longhouses, partible paternity (cf. Lizarralde & Lizarralde 1991; Beckerman & Valentine 2002), and the existence of only few non-compound number terms (cf. Rincón & Quesada 2001/02: 16). In contrast, the Muisca had a system of nine non-compound number terms and, as an 'Andean' feature, as it were, metallurgy (Steward 1948: 9). In terms of cultural heterogeneity of Chibchan-speaking groups, one may also note that Muisca goldwork (10^{th} to mid-16^{th} century A.D., present-day department of Cundinamarca, Colombia) is sometimes remarkably different from Tairona goldwork, manufactured in the Sierra Nevada de Santa Marta, likewise Colombia (10^{h} to mid-16^{th} century A.D.) (cf. Benson 1985: 166–7, 170–1). Differences have also been observed between Muisca pottery and

that from the Isthmian–Caribbean area (Bray 1997: 43). Finally, Chibchan peoples from northern South America, such as the Kogi, Tunebo, and Barí, have a semi-sedentary settlement pattern, a phenomenon which has not been observed among Chibchan peoples of Central America, such as the Bribri and Cabécar (Stephen Beckerman, p.c.).

Some of these heterogeneities in cultural terms may be the result of relatively recent developments. Considering earlier periods, there is a remarkable overlap between the distribution of Chibchan languages and the spread, after A.D. 600, of a particular and uniform gold working style in the Intermediate Area, manufactured in the co-called 'international style' (Bray 1997: 38–9). In Panama for instance, this 'international style' goldwork (ca. A.D. 400 to 900) followed a phase of goldwork in the 'initial style' (before A.D. 500), which had probably been introduced from Colombia (Bray 1992: 34). The metalwork of the 'international style' covers a range of figures that are very typical for the Chibchan area, however, apparently not including the regions were Rama and Pech are spoken at present (see figure 2).

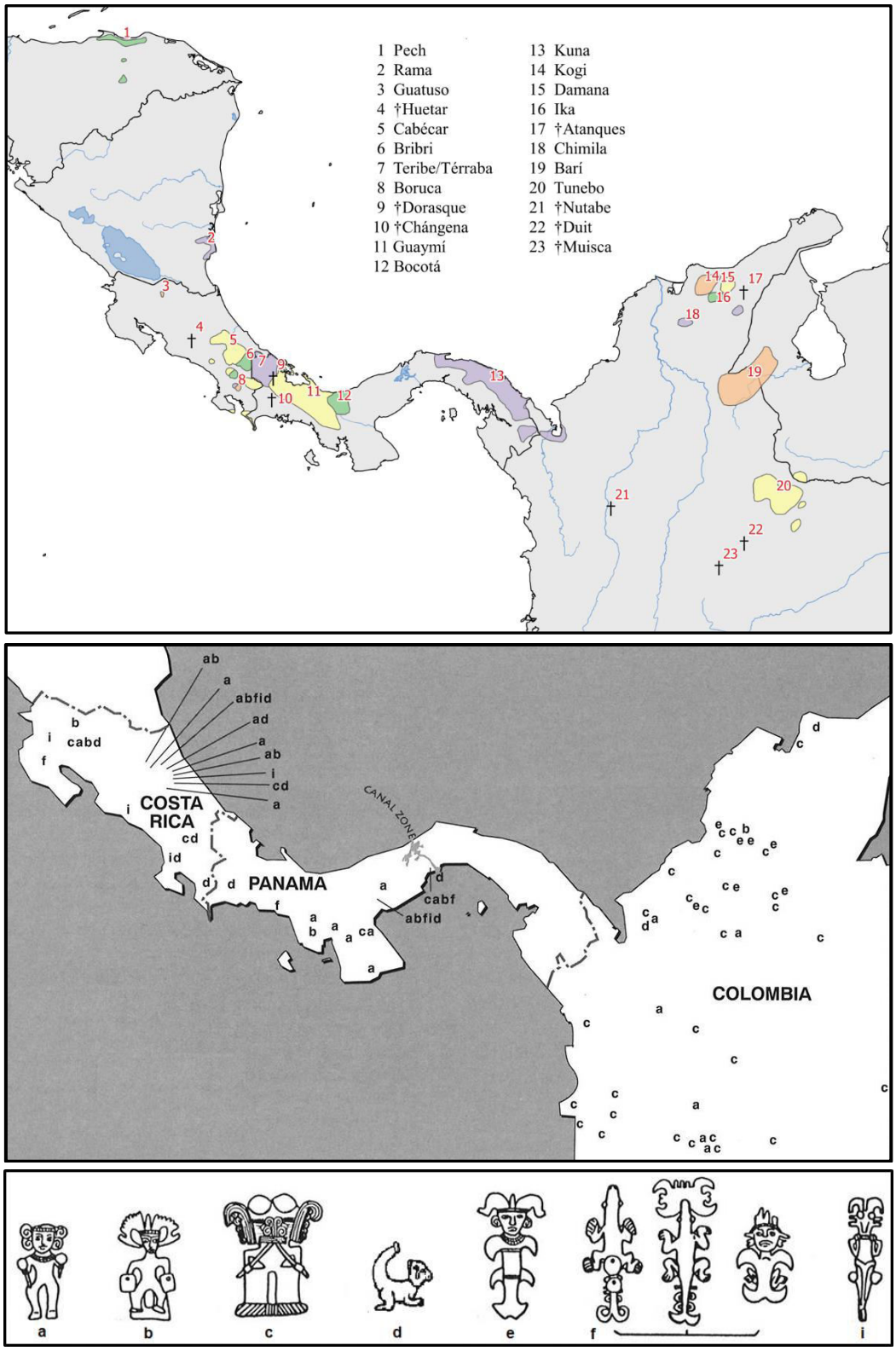


Figure 2: Distribution of Chibchan languages (map from figure 1) and of 'international style' artifacts (A.D. 400–900) (Bray 1992: 38).

Archaeological observations like these may eventually turn out to be relevant in our understanding of the spread of Chibchan languages within the Intermediate Area, given that the distribution of a particular style and iconography in a specific region can be a correlate of linguistic expansion (e.g., Beresford-Jones & Heggarty 2010; Adelaar 2012a: 467).[1]

Before the period of goldwork, stone was predominant in the fabrication of prestige objects throughout the Intermediate Area (predominantly greenstone in the area of present-day Costa Rica, Clados 2017). The carving of jade began around 300 B.C. in what corresponds to present-day Costa Rica (Snarskis 2003: 168–9), and the shift from jade to gold has been located at between A.D. 400 and 700 in this area (ibid.: 175). For an intermediate period, gold and jade objects are still attested side by side at some sites, both in Costa Rica and in Colombia (e.g., in Nahuange, in the Sierra Nevada de Santa Marta) (Bray 2003: 330), and in certain cases the same objects were manufactured in jade and in gold (Snarskis 1998: 89). Nonetheless, the change in the fabrication of prestige objects in the Intermediate Area, from greenstone to gold, has been described as quite clear-cut, dramatic and fundamental (Quilter 2003: 8). The possible correlates of these and other cultural changes at that time, in terms of prehistoric migrations in the Intermediate Area, are not yet fully understood.

Some cultural features and archaeological findings suggest the existence of connections between indigenous populations from the Intermediate Area and South America. It goes without saying that it is difficult to claim that the culture of Central American peoples such as Chibchan-speaking groups is 'South American', given the cultural heterogeneity attested in this latter part of the world – however, more specific cultural parallels can be identified between the Chibchan-speaking and other Central American peoples on the one hand and South American populations on the other hand than with North American peoples (Kirchhoff 1943). Kidder (1940: 444) observes that Central American "Chibchan groups […] had fundamentally South American cultures", and Mason (1940: 76) concurs in that "[t]he Middle American Chibchan tribes are all basically South American in culture." Cultural similarities of peoples from the isthmus with South American peoples have also been mentioned by other authors, mostly in the first half of the twentieth century (e.g., Lothrop 1937: 202; Steward 1948: 10; Stone 1948a: 192–3; Palmatary 1950: 346–7; Snarskis 2003: 188–9), in particular for

[1] The factors triggering the spread of Chibchan-speaking groups in the Intermediate Area remain to be identified. Jolkesky (2016: 230) hypothesizes that cataclysms such as the eruption of the Barú in Panama (Chiriquí Province) at about A.D. 600 (cf. Linares et al. 1975) might have been a driving force behind migratory movements of Chibchan-speaking groups within the Intermediate Area. Constenla Umaña (1981: 353) suggests that the southward migration of Mesoamerican peoples (the Mangue were probably the first of them) and their occupation of Central American territories since about A.D. 600 may have determined some Chibchan migrations.

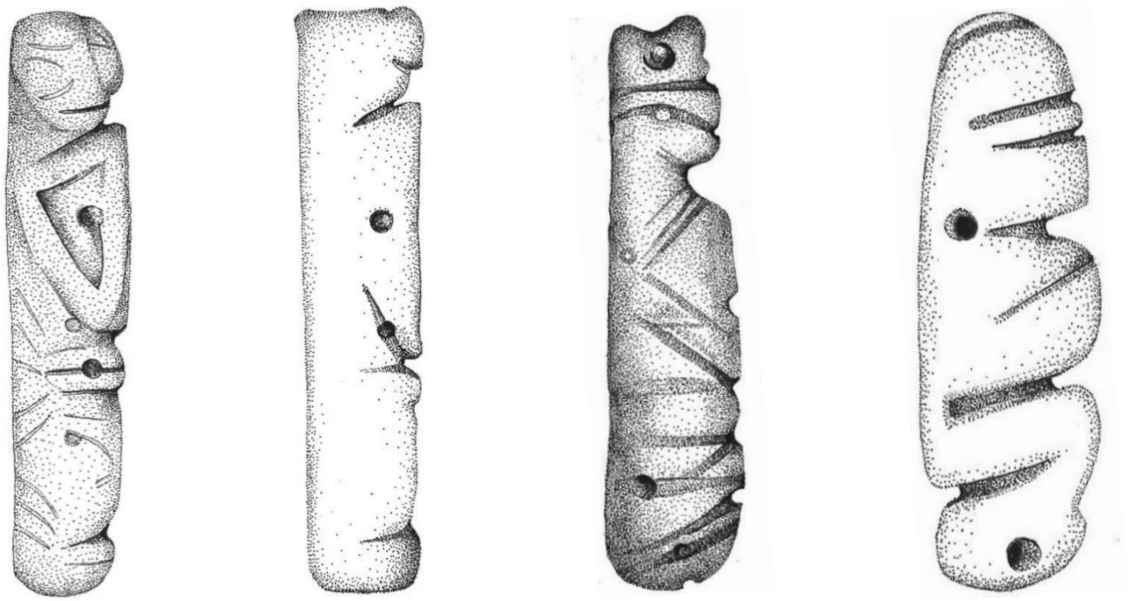
populations of the Atlantic side of the isthmus (e.g., Lehmann 1920: 154). These similarities concern the domains of architecture, ceramics, and prestige goods.

Metallurgy in the Intermediate Area reflects influences from the western part of South America (e.g., Steward 1948: 9) while South American connections are also reflected in the domain of greenstone work. First, it has been known for a long time that greenstone objects called *muiraquitãs*, mostly representing batrachians, were common in the areas of the Tapajós and of the Trombetas rivers in eastern Brazil (southern and northern tributaries, respectively, of the Lower Amazon) (Palmatary 1960: 75–87, 230–5; La Condamine 1745: 140; Costa et al. 2002). Similar objects were probably also looted on the Atlantic watershed of Costa Rica (cf. Mora-Marín 2015: 28) and were known among the Kogi of the Sierra Nevada de Santa Marta (Wassén 1934: 337; a specimen from this area is exhibited in the *Bonner Altamerika-Sammlung*, University of Bonn).[2]

A hitherto unnoticed parallel between greenstone artifacts from the Intermediate Area and a greenstone artifact from the lower Amazon area is shown in figure 3. It depicts some pendants found

1. in a tomb in the Tairona area (Nahuange (figure 3a) and Pueblito (3b) sites in the Sierra Nevada de Santa Marta, Colombia, first published in Mason 1936, plate xcii);
2. in Costa Rica (figure 3c, artifact from the *Bonner Altamerika-Sammlung*, University of Bonn). Other, similar pendants were found in the La Fortuna District, San Carlos Canton, Alajuela Province, central-north Costa Rica (cf. Bray 2003: 330–1);
3. in the Trombetas area (figure 3d) (Palmatary 1960: 85, 233). As stated above, the Trombetas river is a northern tributary of the Lower Amazon of eastern Brazil (state of Pará).

[2] *Muiraquitãs* have also been found in other parts of eastern South America (Wassén 1934), for instance in Surinam (de Goeje 1931), and at archaeological sites southeast of Lake Maracaibo and in the area of Lake Valencia, both in Venezuela (Wassén 1934: 343). Wassén (ibid.: 336–7) observes that frogs made of greenstone are mentioned as a bride-price both among indigenous populations of the state of Maranhão (Brazil) (Heriarte [1662] 1874: 19) and in some oral traditions of the Kogi (Preuss 1926: 159).



(a) **(b)** **(c)** **(d)**

Figure 3: Stone figure pendants (a) from the Nahuange site 1 tomb and (b) from Pueblito site 31 (Mason 1936: plate xcii); (c) Costa Rican greenstone-pendant from the Bonner Altamerika-Sammlung; (d) greenstone pendant, found in the Trombetas area (state of Pará, Brazil) (Palmatary 1960: 85, 233). The artifacts were drawn by Maria Magdalena Antczak.[3]

The Colombian pendants depicted in figure 3a and 3b are the vertical halves of split jade figures representing humanoids. Only one half of each figure was deposited in the tomb (Bray 2003: 330–1). As to the age of these artifacts, the pendant in figure 3a was found in the same tomb as a golden figurine dated A.D. 310 +/- 70 (Bray 2003: 324). The pendant, depicted in figure 3d, from the Trombetas area may have been elaborated by the bearers of the Konduri culture, which has some similarities with and a similar time depth as the Santarém/Tapajós culture of the lower Amazon (cf. Guapindaia 2008: 38) (Early work in Santarém ceramics dates from A.D. 900 (Gomes 2001: 138).) More, somewhat similar stone pendants have been found in the Lower Amazon area (cf. Palmatary 1960: 231–7; de Goeje 1931: 452). Whereas the split-figure pendants from the Intermediate Area seem to have represented split humanoids (figures 3a–c), the pendant in figure 3d seems to have represented a split batrachian. The artifacts shown in figure 3 are similar in terms of material (greenstone in the case of 3a, 3c, and 3d), shape (two-dimensional), and in the arrangement of incisions and holes. Additionally, the fact that they are all split figures suggests that a certain similarity in belief systems existed

[3] The pendants measure: a) 13,6x2,8x0,8cm (Mason 1936: 191); b) 20,5x4,3x1cm (Mason 1936: 191); c) 9,8x2,1/1,4x1,1/0,1cm; and d) 4,5 cm (Palmatary 1960: 233).

between the populations from the Intermediate Area and the Lower Amazon during a certain period of time. It would be difficult to claim that split-figure pendants in the Lower Amazon area, such as the one depicted in figure 3d, were recent importations from the Intermediate Area, since in this case, one would also expect gold artifacts from this region to have reached the lower Amazon area, which does not seem to be the case.[4] In sum, there are some obvious parallels in the greenstone work of the Intermediate Area and of the Lower Amazon; the origin of these parallels is unknown and more research is necessary to determine the relationship. Notwithstanding, the observations made here suggest a potential connection of Chibchan-speaking groups or their ancestors with peoples from the eastern Lowlands of South America.

1.1.3 Genetics

In terms of genetic features, populations speaking Chibchan languages show a rather homogeneous picture and have peculiarities that set them apart from other indigenous groups in the Americas (e.g., Barrantes 1990; Yunis et al. 2013). It has repeatedly been observed, for instance, that the diversity of mitochondrial DNA (mtDNA) is relatively low among populations speaking Chibchan languages (Melton et al. 2007). Four main mtDNA haplogroups occur in the Americas (A, B, C, and D), indicating four founding mother lineages of Amerind people (Schurr & Sherry 2004; Merriwether et al. 1995), and a clear pattern emerges for peoples speaking Chibchan languages, namely a high proportion of haplogroup A, and a near absence of haplogroup D (Torroni et al. 1994; Batista et al. 1995; Keyeux et al. 2002; Santos et al. 1994; Melton et al. 2007; Melton 2008). A possible interpretation of this relatively uniform picture is genetic drift, due to a small founding population (cf. Mayr 1942) with little subsequent gene flow, that is, little subsequent admixture of women with different mtDNA. The investigation of autosomes, which include information from both parents, likewise produces a homogeneous picture: A slightly lower level of heterozygosity among Chibchan-speaking groups (compared with neighboring populations) was observed by Wang et al. (2007: 2060), who investigated genome-wide patterns of variation among Amerindian peoples, namely 678 autosomal microsatellite markers. These reduced levels of heterozygosity among Chibchan-speaking groups may likewise reflect genetic drift, that is, a rather small Proto-Chibchan founding population, which may have spread relatively recently or not over the Intermediate Area.

[4] The only gold artifact from the Intermediate Area that arguably reached eastern South America is a double-headed eagle or vulture made of *guanín*, which was dredged from the bottom of the mid-Mazaruni river in Guyana (Whitehead 1990), and which is “pure Colombian in style” (Bray 1997: 47).

Genetic data suggest a relatively close link between Chibchan-speaking groups and peoples from eastern South America. In terms of external relations as reflected in autosomal markers, Wang et al. (2007) find strong support (97%) for clustering together the Cabécar, Guaymí (Ngäbe), Ika, Kogi, with the geographically adjacent Wayuu (Maipuran), Zenú (unknown linguistic affiliation), Emberá and Waunana (both Chocoan). Additionally, the authors propose "a reasonably well-supported cluster (86%)" including all South American populations outside the Andes (ibid.: 2054), that is, the eight aforementioned groups including all considered Chibchan populations, the Aché (Tupí–Guaraní), Guaraní (Tupí–Guaraní), Inga (Quechuan), Karitiana (Tupían), Piapoco (Maipuran), Surui (Tupí–Guaraní), Ticuna (Ticuna–Yurí), and Kaingáng (Southern Jê) (ibid.: 2057). This suggests that in terms of autosomal markers, Chibchan groups are more closely related to populations from eastern South America than to Mesoamerican populations speaking Mixe–Zoquean or Mayan languages. The two latter groups cluster more closely together with Andean populations (Wang et al. 2007: 2060). Genetic evidence would thus be in line with an original homeland of Chibchan in eastern South America.

1.2 Chibchan historical linguistics

In this subsection, I will briefly introduce Chibchan historical linguistics, presenting, first, the discovery of Chibchan as a genealogical unit (1.2.1), followed by some information about the internal structure of the family (1.2.2) and external classifications proposed so far (1.2.3).

1.2.1 Establishment of the family and reconstructive work

The existence of the Chibchan language family as we know it today (see figure 1 above) became known to western researchers only relatively recently, namely since the end of the nineteenth century. At that time, the existence of different South American families, such as Maipuran, Cariban, Guaicuruan, Sáliba–Piaroan, Guahiboan, and Tupí–Guaraní had by and large been recognized (cf. Gumilla 1741: 316; Gilij 1782: 202, 286, 345–6, 390–2; Pache et al. 2017). Herzog (1886) was the first author who clearly treated certain Chibchan languages as genealogically related. Before him, some authors might have suspected a genealogical relationship between Chibchan languages, but none of them made explicit statements.[5] In his pioneering work, Herzog (1886) took for granted rather than argued for a genealogical link between the following Chibchan languages of Central America: Boruca, Bribri, Cabécar, Guatuso, and Teribe. He provided evidence for a genealogical connection of these languages with Kuna and observed some similarities with Muisca. Soon after this publication, different researchers contributed to the further establishment of the family:

- A genealogical connection between the Colombian Chibchan languages Kogi, Ika, Damana and Muisca was first argued for by Müller (1888: 189). A similar link had tentatively been ventured before by Sievers (1886: 400).[6]
- The best-known and most influential early paper on Chibchan linguistic prehistory and genealogical unity is Uhle's (1890) work. It was first presented in 1888, at the Congress of Americanists. Uhle must be credited with establishing the relatedness of Chibchan languages via recurrent sound correspondences. Also, he was the first author to definitely connect Chibchan languages from South America (Damana, Ika, Kogi, Muisca) with their sister languages from Central America (Bocotá, Boruca, Bribri, Cabécar, Guaymí, Teribe/Térraba).

[5] Piedrahita (1688: 142), may have considered a particular connection to exist between Muisca and an indigenous language of the Sierra Nevada de Santa Marta (cf. Uhle 1890: 468). Instead, Cassani (1741: 48) seems to have considered Tunebo, including Sínsiga (present-day Boyacá department), as a dialect of Muisca (cf. Uricoechea 1871: xxxiv). Finally, Thiel (1882: 155) seems to have suspected a particular link to exist between extinct Huetar and Cabécar, two Chibchan languages of Costa Rica.

[6] According to Sievers (1886: 400), Ika (Arhuaco) can probably be connected to Muisca: "[…] vielleicht neigt sie [i.e. Ika [MP]) zum Chibcha. (?)"

Finally, Uhle (1890: 485) was also the first to stress the existence of some lexical similarities with Chimila.

- One year later, Brinton corroborated the Chibchan status of Chimila (1891a: 183, 345). He also connected Rama and Chánguena with each other (ibid.: 367), identified Duit, Sínsiga and Tunebo as Chibchan languages, and seems to be the first author to refer to the family as 'Chibcha' (cf. ibid.: 181–9).

Most authors, after Uhle's (1890) paper, tried to link particular Chibchan languages with the family as a whole. Table 2 illustrates further explicit proposals for a Chibchan status of languages others than those mentioned above. A similar overview is given by Kaufman (1988).

TABLE 2
EARLY WORKS EXPLICITLY PRESENTING CERTAIN LANGUAGES AS CHIBCHAN

Language presented as Chibchan	First publication
Huetar	Brinton (1897); Lehmann (1920: 236–7)
Dorasque, Kuna	La Grasserie (1904)
Nutabe, Catío	Rivet (1943a)
Barí	Rivet and Armellada (1950)
Pech	Holt (1975), cited in Campbell (1979: 942–4)

In sum, as an undeniable genealogical unit, Chibchan was discovered piecemeal over a time span of about a hundred years, between 1886 and 1975 (cf. Herzog 1886; Holt 1975, cited in Campbell 1979: 942–4).

Reconstructed, asterisked forms were first proposed in the 1940's by Holmer (1947: 207–8), based on evidence from Muisca, Bribri, Rama, Kogi and Kuna. The validity of Chibchan as containing the languages in the previous paragraphs was finally confirmed in the studies of Wheeler (1972), Levinsohn (1975), Kaufman (1988), and above all in the works of Constenla Umaña (1981) and Holt (1986), all providing regular sound correspondences and/or extensive sets of reconstructed forms. Table 3 juxtaposes some reconstructions of Constenla Umaña (1981) and Holt (1986) which are quite different in some cases.

TABLE 3
SOME DIVERGENT PROTO-CHIBCHAN RECONSTRUCTIONS PROPOSED BY CONSTENLA UMAÑA (1981) AND HOLT (1986)

English	Proto-Chibchan (Constenla Umana 1981)	Proto-Chibchan (Holt 1986)
'arm, hand'	*kuda 'arm'	*guLÀ 'hand'
'excrement'	*ˈgə́	*ša 'excrement, defecate'
'nose'	*dəˈIkI	*dɨk' ~ *dik'
'take'	*ˈgúʔ	*ku 'get, take, receive'
'water'	*ˈdíʔ	*di

1.2.2 Internal structure of the family

In terms of the internal structure of the Chibchan family, the earliest proposal was made by Uhle (1890: 474–5), although he did not explicitly refer to the concept of shared innovations as a basis of subgrouping Chibchan languages. The units proposed by Uhle (1890) are (1) Muisca; (2) Damana, Ika, Kogi; (3) Bribri, Cabécar, Teribe, Térraba, Boruca; (4) Guaymí, Bocotá. Since then, subgrouping matters have been followed up by several authors, for instance by Holmer (1947: 205–7), Kaufman (1988), and above all by Constenla Umaña (e.g. 1981, 2008, 2012). Constenla Umaña's investigations resulted in a tree model which is briefly outlined below. Such a tree model also provides some non-linguistic information about the speakers' prehistory, since nodes can be interpreted as diversification events, possibly implying migration (François 2015: 163–5). In the case of Chibchan, Constenla Umana (2012: 416–7) proposed a fundamental split of the Chibchan family into Pech on the one hand, and 'Core Chibchan' on the other hand. Pech is the most northwestern member of Chibchan (see figure 1). The internal structure of the Chibchan family, according to Constenla Umaña's (2012) proposal, is as follows:

I Pech
II Core Chibchan
    IIA Votic: Rama, Guatuso
    IIB Isthmic
        B1 Western Isthmic
            B1.1 Viceitic: Cabécar, Bribri
            B1.2 Teribe/Térraba
            B1.3 Boruca
        B2 Doracic: Dorasque
        B3 Eastern Isthmic
            B3.1 Guaymiic: Guaymí, Bocotá
            B3.2 Kuna
    IIC Magdalenic
        C1 Southern Magdalenic
            C1.1 Chibcha: Muisca, Duit
            C1.2 Tunebo
            C1.3 Barí
        C2 Northern Magdalenic
            C2.1 Arhuacic
                C.2.1.1 Kogi
                C.2.1.2 Eastern-southern Arhuacic
                    C.2.1.2.1 Eastern Arhuacic: Damana, Atanques
                    C.2.1.2.2 Ika
            C2.2 Chimila

Constenla Umaña's (2012) classificatory work shows that there is a more fundamental diversity among the Chibchan languages of Central America than among the 'Magdalenic' Chibchan languages of northern Colombia. This is also in line with the findings of Wichmann et al. 2010, which are based on a quantitative approach. This higher diversity of Chibchan languages in Central America suggests that the presence of these Chibchan-speaking groups in Central America is relatively old, whereas the presence of Chibchan-speaking groups in northern South America is the result of a relatively recent, eastward migration. The scenario behind this argument is the following: Language diversity takes some time to develop (de Laet 1643, cited in Wright 1917: 268; Jefferson 1984: 227, cited in Campbell 1997: 36). From a homeland with some existing linguistic diversity, a homogenous group of people separates, all speaking the same language. Later, in the newly populated territory, the formerly homogenous group begins to split as well. These splits, however, are less fundamental than the diversification in the homeland, since they are more recent (Sapir 1916; Dyen 1956; Constenla Umaña 1981: 352–3). The area with the most fundamental diversity (in this case, Central America) can therefore be considered as the center of dispersal of the proto-language in question, Proto-Chibchan in this case.

In other instances, the internal classification of Chibchan and its interpretation in terms of prehistory still leave some questions open. As usual, for example, there is some criss-crossing of isoglosses (Constenla Umana 2012: 413). Which isoglosses reflect joint migrations and which are due to later contact is often difficult to decide (cf. François 2015: 167–72).

Also, the status of Pech as relatively isolate within Chibchan seems to be debatable, given that it shares several features with languages that belong to the subgroup of Magdalenic (Colombian) Chibchan languages. Some of the parallels in question are shown in table 4; they include, among others, shared innovations in phonology, variation or phonotactics (1–3), morphology (4–8), and lexicon (9).

TABLE 4
SOME INNOVATIONS SHARED BY PECH AND MAGDALENIC LANGUAGES

	Pech	Magdalenic Chibchan languages
1	*ʦ > *t* / __*u* (cf. subsection 2.3.20 below)	*ʦ > *t* / __*u* in Atanques, Barí, Damana, Ika (cf. subsection 2.3 below)
2	*jìwi ~ jùwi* ‘moon, month’ (Holt 1986: 100)	Kogi <nyui ~ nui> ‘sun’ (Preuss 1927: 521)
3	A sequence *nɾ* is frequently broken up by *d*: *arà꞉-n-ɾiʔ* ‘they made a noise’, for instance, is realized as [arà꞉ndɾiʔ] (Holt 1999a: 25)	The Tunebo sequence *nr* is realized as a soft [ndr] (Headland 1997: 9).
4	*pi-…-wá* discontinuous second-person plural object marker (Holt 1999a: 42)	*mə-…-ua* discontinuous second-person plural marker in Damana (Trillos Amaya 1999: 43)
5	*-píɾ/-páɾ* preinceptive (Holt 1999a: 48)	*-pən ~ -pan* inchoative in Ika (Frank 1985: 19)[7]
6	*-ti* future possibility, *-nih* speculative (Holt 1999a: 58)	*-ˈni* conjecture, inference in Damana (Trillos Amaya 1999: 48)
7	*-t ~ -n* durative, iterative (Holt 1999a: 27)	*-ən* imperfective in Ika and Damana (Frank 1985: 6; Trillos Amaya 1999: 44)
8	Imperative suffix < *-u (Constenla Umaña 2012: 414)	Imperative suffix < *-u in Damana, Ika, Kogi, Muisca, Tunebo (Constenla Umaña 2012: 414)

[7] The variation in Pech and Ika probably reflects independent processes. It is driven by semantic considerations – person marking – in Pech (Holt 1999a: 48), and probably by allophonic variation in Ika.

9	'10' < *uka (cf. subsection 2.2 below)	'10' < *uka in Damana, Ika, Kogi, and Tunebo (Constenla Umaña 2012: 417)

Eventually, shared innovations such as those shown in table 4 may turn out to have fundamental implications for the internal structure of the Chibchan family. In any case, they suggest the possibility of a southeastern homeland of Pech, and a relatively recent migration of this group to the northwest.

1.2.3 External classification

The question of Chibchan external classification of Chibchan has been discussed since the end of the nineteenth century (cf. Herzog 1886), and in the twentieth century, Chibchan languages have often been used as the basis for a kind of 'ragbag' in the classification of Native American languages (Quesada 2007: 17). Indeed, Chibchan has probably been connected to more other Native American languages than any other family: Different researchers proposed large genealogical units, such as 'Macro-Chibcha(n)' (Jijón y Caamaño 1943; Mason 1950; Greenberg 1960, cited in Key 1979: 29; Key 1979), 'Chibchan–Paezan' (Greenberg 1987) or a 'Chibcha stock' (Loukotka 1968). Their proposals were largely based on multilateral comparison and 'look alikes', not on recurrent sound correspondences. Table 5 shows the languages that have been included in these macro-proposals, together with their areas of distribution and the authors that classified these languages as 'Macro-Chibchan'. A detailed account on previous proposals can be found in Kaufman (1988).

TABLE 5
LANGUAGES THAT HAVE BEEN CONNECTED TO CHIBCHAN IN THE CONTEXT OF MACRO-CHIBCHAN AND SIMILAR PROPOSALS

Language	Area of distribution	Author including it into his Macro-Chibchan stock
Andakí (unclassified)	Southwestern Colombia	Ria; Lo; Gr ("Nuclear Paezan")
Barbacoan	Southwestern Colombia, northwestern Ecuador	Sch; Ria; Lo; Gr ("Nuclear Paezan")
Betoi (Sáliba–Piaroan, according to Zamponi 2014)	Western Venezuela	Ria; Lo; Gr ("Chibchan–Paezan")
Candoshi–Shapra (unclassified)	Northern Peru	JC

Cariban	Northeastern South America, central Brazil	Her; Sch
Chocoan	Western Colombia, western Ecuador, eastern Panama	JC; Gr (“Nuclear Paezan”)
Cholón (Cholón–Hibito)	Central Peru	JC
Cofán (unclassified)	Northeastern Ecuador, southwestern Colombia	JC
Cuitlatec (unclassified)	Southwestern Mexico	Gr (“Chibchan”); Swa (1967)
Esmeraldeño (unclassified)	Northwestern Ecuador	Sch, JC; Lo
Huarpean	Western Argentina	Gr (“Chibchan–Paezan”)
Itonama	Northeastern Bolivia	Gr (“Chibchan–Paezan”)
Jicaquean	Central Honduras	JC
Jirajaran	Western Venezuela	Lo; Gr (“Chibchan–Paezan”)
Kamsá (unclassified)	Southern Colombia	Lo
Kunza (unclassified)	Northern Chile	Gr (“Chibchan–Paezan”)
Lencan	Eastern El Salvador, southern Honduras	JC; Gr (“Chibchan”)
Maipuran	Western, central and northern South America	Sch
Mayan	Southern Mexico, Guatemala, Belize, western El Salvador and western Honduras	Sch
Misumalpan	Area extending from (north-)eastern and central Nicaragua to eastern San Salvador	JC; Lo; Swa; Gr (“Nuclear Chibchan”)
Mochica (unclassified)	Northwestern Peru	JC; Gr (“Chibchan–Paezan”)
Mosetén	Northwestern Bolivia	Sch
Mura (Mura–Pirahã)	Northwestern Brazil	Gr (“Chibchan–Paezan”)
Páez (unclassified)	Southwestern Colombia	Sch, Ria; JC; Lo; Gr (“Chibchan–Paezan”)
Panoan	Eastern Peru, western Brazil, northern Bolivia	Swb

Pano–Tacanan	Eastern Peru, western Brazil, northern Bolivia	Sch; Ho
Pijao (possibly Cariban, cf. Constenla Umaña (1991: 62))	Southwestern Colombia	Ria
Pumé (probably related to Chocoan, see subsection 3.2.3.2 below)	Central Venezuela	Sch, JC; Lo
Purépecha (unclassified)	Western Mexico	Gr ("Chibchan")
Subtiaba (Tlapanec–Manguean)	Western Nicaragua	JC
Timote (Timote–Cuica)	Northwestern Venezuela	Sch, JC
Timucua (unclassified)	Northern and central Florida, southeastern Georgia	Gr ("Chibchan–Paezan")
Tucanoan	Southern Colombia and adjacent regions in Ecuador, Peru and Brazil	Swa
Tupí–Guaraní	Large parts of Lowland South America	Her
Uto–Aztecan	Southwestern North America	Swa, Swb; Ho
Warao (unclassified)	Coastal regions of northeastern Venezuela, Guyana, Suriname	Her; Gr ("Chibchan–Paezan")
Yurumanguí (unclassified)	Western Colombia	JC
Xinkan	Southeastern Guatemala and adjacent parts of El Salvador and Honduras	JC; Gr ("Chibchan")
Yanomaman	Southern Venezuela, northwestern Brazil	Gr ("Chibchan")
Yurí (Ticuna–Yurí)	Northwestern Brazil, southeastern Colombia	Sch

Sources are Herzog (1886); Schuller (1919/20); Rivet (1924a: 680–4); Jijón y Caamaño (1943: 419–20); Swadesh (1954, 1967); Loukotka (1968: 233–53); Holt (1986, 1988); Greenberg (1987: 106, 382).[8]

[8] In terms of subgrouping, Greenberg (1987: 106, 382) makes a distinction between "Chibchan–Paezan", "Chibchan", "Nuclear Chibchan" and "Nuclear Paezan". Lévi-Strauss' (1948) proposal of a Chibchan–Nambikwaran connection is not shown in table 5 since the author interprets the parallels that he observes as the result of contact (ibid.: 192).

None of the sometimes contradictory proposals shown in table 5 has met with favorable judgements in the works of Constenla Umaña (e.g., 1981: 8–14, 36–70, 2012: 418), except Chibchan connections with Tacanan, Chocoan, and, above all, Lencan and Misumalpan (Constenla Umana 1981: 14; 2002, 2005, 2012: 418; Constenla Umaña & Margery Peña 1991: 172). Remarkably, most languages that have been classified as connected with Chibchan, according to table 5, are spoken in South America.

1.3 Structure of this thesis

Dealing with Chibchan historical linguistics, this thesis is divided into two parts: The first addresses Chibchan-internal issues and the second addresses the external classification of this language family. It goes without saying that the second part of this thesis is largely based on the results presented and discussed in the first part.

Section 2 deals with Proto-Chibchan and internal issues. It starts with a subsection on methodology (2.1). The different cognate sets and reconstructed Proto-Chibchan forms are provided in subsection 2.2. This is followed by a detailed discussion of the reflexes of Proto-Chibchan phonemes that are attested in the different Chibchan languages (2.3). Subsection 2.4 presents and discusses Proto-Chibchan phonemes, suprasegmentals, phoneme alternation and subphonemic variation, and phonotactics. Section 2 closes with a summary and discussion of Proto-Chibchan as reconstructed here (2.5).

Section 3 deals with the external classification of Chibchan. This section begins with a discussion of methodological issues in the investigation of distant genealogical relationships (3.1). This is followed by a subsection presenting previously proposed external genealogical connections of Chibchan in some detail (3.2). An alternative proposal is discussed in subsection 3.3: Evidence for a genealogical link between Chibchan and the Macro-Jê stock of eastern South America. Section 3 closes with a conclusion and discussion (3.4).