Arawak Languages

Chapter · December 2006
DOI: 10.1016/B0-08-044854-2/02292-6

1 author:

Some of the authors of this publication are also working on these related projects:

Serial verbs View project

is the religious poetry of Saint Ephrem, which was much admired and imitated even beyond the Syriac language area. From the 10th century, Arabic replaced Syriac among Christians as the chief language of theology, philosophy, and medicine, but the 13th century saw a veritable West Syriac renaissance, embodied especially in the great polymath Bar Hebraeus, who wrote with equal facility in Syriac and Arabic. In contrast to the wide use of Syriac, Syro-Palestinian Christian Aramaic (alternatively designated Syro-Palestinian Syriac because it was written in the West Syriac script) was employed only in Palestine and Syria, and the extant texts (mostly biblical, liturgical, or hagiographical) are all translations from Greek.

Spoken Aramaic dialects have been in continuous use in a number of places right into modern times. Modern Western dialects of Aramaic are spoken, by Christians and Muslims, in three villages north of Damascus, namely Ma'lula, Bah'a, and Jubb 'Addin. Eastern dialects have been more extensively used by Christians in various localities. In the mountainous area of Southeast Turkey known as Tur 'Abdin, Turoyo ('the mountain language') is spoken by members of the Syrian Orthodox Church. Other Eastern Aramaic dialects have been spoken in modern times by the Jews of Kurdistan and Azerbaijan, most of whom have now emigrated to Israel, and a modern Mandaic dialect has survived in Iran. The greatest use of Aramaic in modern times, however, has been by East Syrian Christians, among whom a number of East Aramaic dialects have been employed. Modern literary Syriac (Swadaya) may be said to have begun with the printing of books in the local dialect by the American Presbyterian Mission at Urmia in Northwest Iran. Although the number of people currently using some form of Aramaic is small, their determination to keep it alive is a testimony to their pride in a language whose demonstrable lifespan extends to 3000 years.

See also: Ancient Near-Eastern Religions; Arabic; Bible; Bible Translations: Ancient Versions; Christianity in Central Asia and the Near East; Hebrew, Biblical and Medieval; Iran: Language Situation; Iraq: Language Situation; Israel: Language Situation; Judaism; Lebanon: Language Situation; Semitic Languages; Syria: Language Situation; Syriac; Translation: History; Turkey: Language Situation.

Bibliography

Brock S P (1989). 'Three thousand years of Aramaic literature.' *Aram 1(1),* 11–23.

Brock S P (1996). *Syriac studies: a classified bibliography, 1960–1990.* Kaslik, Lebanon: Université Saint-Esprit de Kaslik.

Fitzmyer J A (1979). 'The phases of the Aramaic language.' In Fitzmyer J A (ed.) *A wandering Aramean: collected Aramaic essays.* Missoula, MT: Scholars Press. 57–84.

Fitzmyer J A & Kaufman S A (1992). *An Aramaic bibliography.* Baltimore and London: Johns Hopkins University Press.

Aranda   *See:* Arrernte.

Araucanian   *See:* Mapudungan.

Arawak Languages

A Y Aikhenvald, Institute for Advanced Study, La Trobe University, Bundoora, Victoria, Australia

The Arawak language family contains the largest number of languages in Latin America. Geographically, it spans four countries of Central America – Belize, Honduras, Guatemala, Nicaragua – and eight of South America – Bolivia, Guyana, French Guiana, Surinam, Venezuela, Colombia, Peru, Brazil (and also formerly Argentina and Paraguay).

There are about 40 living Arawak languages. The first Native American peoples encountered by Columbus – in the Bahamas, Hispaniola, and Puerto

Rico – were the Arawak-speaking Taino. Their language became extinct within a hundred years of the invasion. Spanish and many other European languages inherited a number of loans from Arawak languages. These include widely used words such as *hammock*, *tobacco*, *potato*, *guava,* and many other names for flora and fauna.

The creation of a mixed language of Arawak/Carib origin in the Lesser Antilles is one of the most interesting pieces of evidence on language history in pre-conquest times. Speakers of Iñeri, a dialect of the Arawak language now (misleadingly) called Island Carib, were conquered by Carib speakers. They developed a mixed Carib/Arawak pidgin that survived until the 17th century (Hoff, 1994). Speech of men and speech of women were distinguished in the following way. Women used morphemes and lexemes of Arawak origin, while men used lexical items of Carib origin and grammatical morphemes mostly of Arawak origin. The pidgin coexisted with Carib used by men and Iñeri used by women and children; it belonged to both parties and served as a bridge between them. This diglossia gradually died out with the spread of competence in Island Carib among both men and women. As a result, Island Carib, an Arawak language, underwent strong lexical and, possibly, grammatical influence from Carib.

The languages in areas settled by the European invaders soon became extinct. Those on the north coast of South America perished first, before 1700. When the search for gold and rubber extended up the Amazon and its tributary the Rio Negro, further languages succumbed, from the 18th century up until the present day. Sometimes the Indians retaliated, attacking settlements and missions; but the invaders always returned. Indian rebellions often provoked forced migrations which sometimes ended up creating a new dialect or even a language. For instance, in 1797 the British authorities removed the rebellious inhabitants of St. Vincent (an island in the Lesser Antilles) to Belize on the mainland. These were racially a mixture of black slaves and Indians, who spoke Island Carib. This resulted in the creation of a new dialect of Island Carib – known as Central American Island Carib, Kariff, Black Carib, or Garifuna – which by the 20th century had developed into a separate language, now spoken in Central America (Taylor, 1977).

The overwhelming majority of Arawak languages are endangered. Even in the few communities with more than 1000 speakers, a national language (Portuguese or Spanish) or a local lingua franca (Lingua Geral Amazônica, Quechua, or Tucano) is gaining ground among younger people. The few healthy Arawak languages are Guajiro in Venezuela and Colombia (estimates vary from 60 000 to 300 000 speakers) and the Campa languages (total estimate 40 000 to 50 000 speakers), one of the largest indigenous groups in Peru.

Most of the materials on Arawak languages collected during the second half of the 20th century are by missionary linguists. Their quality and quantity varies. Only three or four languages have full descriptions available.

The genetic unity of Arawak languages was first recognized by Father Gilij as early as 1783. The recognition of the family was based on a comparison of pronominal cross-referencing prefixes in Maipure, an extinct language from the Orinoco Valley, and in Moxo from Bolivia. Gilij named the family Maipure. Later, it was renamed Arawak by Daniel Brinton after one of the most important languages of the family, Arawak (or Lokono), spoken in the Guianas. This name gained wide acceptance during the following decades. The majority of Native South American scholars use the name Arawak (Aruák) to refer to the group of unquestionably related languages easily recognizable by pronominal prefixes such as *nu-* or *ta-* 'first person singular', *(p)i-* 'second person singular', prefix *ka-* meaning 'have', and negator *ma-*. A number of scholars, mainly North Americans, prefer to use the term Arawak(-an) to refer to a much more doubtful higher-level grouping, and reserve the term Maipuran (or Maipurean) for the group of undoubtedly related languages that are claimed to be one branch of Arawakan (see Payne, 1991). Here I follow the South American practice and use the name *Arawak* for the family of definitely related languages.

The limits of the family were established by the early 20th century. Problems still exist concerning internal genetic relationships within the family and possible genetic relationships with other groups. Reconstruction, internal classification, and subgrouping of Arawak languages remain matters of debate; further detailed work is needed on both the descriptive and comparative fronts.

The putative studies of Arawakan by Ester Matteson, G. Kingsley Noble, and others are deeply flawed. Unfortunately, these have been adopted as the standard reference for the classification of Arawak languages, especially among some anthropologists, archaeologists, and geneticists, influencing ideas on a putative proto-home and migration routes for proto-Arawakan' – see the criticism in Tovar and De Tovar (1984), Dixon and Aikhenvald (1999: 12–15), and Aikhenvald (1999a).

Little is known about a proto-home for the Arawak family. The linguistic argument in favor of an Arawak proto-home located between the Rio Negro and the Orinoco rivers – or on the Upper Amazon – is based on the fact that there is a higher concentration of structurally divergent languages found in this region.

This area has also been suggested as one of the places where agriculture developed. This is highly suggestive and corroborated by a few mythical traditions of northern origin by Arawak-speaking peoples south of the Amazon. The origin myths of the Tariana, in northwest Amazonia, suggest that they could have come from the north coast of South America.

Arawak languages are complicated in many ways. Words can be differentiated by stress in some languages, such as Baure and Waurá (south of Amazonas), and Tariana, Achagua, and Warekena (north of Amazonas). At least two have tones – Terêna in the South, and Resígaro spoken in the far northeast of Peru.

Each Arawak language has a few prefixes and numerous suffixes. Prefixes are typically monosyllabic, while suffixes can consist of one or more syllables. Roots usually contain two syllables. Prefixes are rather uniform across the family, while suffixes are not. What is a free morpheme in one language can be a grammatical marker in another language; for instance, postpositions become causative markers, and nouns become classifiers. An Apurina noun *maka* means 'clothing' – this is where the word for hammock comes from. In Baniwa of Içana, -*maka* is a classifier for stretchable thin extended objects, e.g., *tsaia* 'skirt' or *dzawiya* 'jaguar's skin', as in *apa-maka* (one-CLASSIFIER:CLOTHING) 'one piece of clothing'.

Most grammatical categories in Arawak languages are verbal. Cases to mark subjects and objects are atypical. Tariana, spoken in northwest Brazil, has developed cases for core grammatical relations to match the pattern in nearby Tucanoan languages (Aikhenvald, 1999b).

Arawak languages spoken south of the Amazon (South Arawak) have a more complex predicate structure than those north of the Amazon (North Arawak). South Arawak languages such as Amuesha or Campa have up to thirty suffix positions. North Arawak languages such as Tariana or Palikur have not more than a dozen suffixes. Suffixes express meanings realized by independent words in familiar Indo-European languages, e.g., 'be about to do something', 'want to do something', 'do late at night', 'do early in the morning', 'do all along the way', 'in vain', 'each other'.

Verbs are typically divided into transitive (e.g., 'hit'), active intransitive (e.g., 'jump') and stative intransitive (e.g., 'be cold'). All Arawak languages share pronominal affixes and personal pronouns. Pronominal suffixes refer to subjects of stative verbs and direct objects. Prefixes are used for subjects of transitive verbs and of intransitive active verbs, and for possessors. That is, most Arawak languages are of active-stative type. For instance, in Baniwa one says *nu-kapa*

Table 1 Pronominal prefixes and suffixes in proto-Arawak

Person	Prefixes		Suffixes	
	Singular	*Plural*	*Singular*	*Plural*
1	*nu-* or *ta-*	*wa-*	*-na, -te*	*-wa*
2	*(p)i-*	*(h)i-*	*-pi*	*-hi*
3nf	*r i-, i-*	*na-*	*-r i, -i*	*-na*
3f	*thu-, ru-*	*na-*	*-thu, -ru, -u*	*-na*
'impersonal'	*pa-*	—	—	—

'I see' and *nu-watsa* 'I jump', but *nu-kapa-ni* 'I see him' and *hape-ni* 'he is cold' (*nu-* refers to 'I' and *-ni* to 'him'). And 'my hand' is *nu-kapi*.

Some languages have lost the pronominal suffixes (and with them the morphological basis for an active-stative system); these include Yawalapiti (Xingú area, Brazil) and Chamicuro (Peru) to the south of the Amazon, and Bare, Resígaro, Maipure, and Tariana to the north. The form of the first person pronoun is *ta-* in the Caribbean (Lokono, Guajiro, Añun, Taino) and *nu-* in other languages. This is the basis for classification of Arawak languages into *Nu-*Arawak and *Ta-*Arawak.

Proto-Arawak must have had an unusual system of four persons: first, second, third, and impersonal. The forms of prefixes and suffixes reconstructed for proto-Arawak are given in Table 1.

Most Arawak languages distinguish two genders – masculine and feminine – in cross-referencing affixes, in personal pronouns, in demonstratives, and in nominalizations, e.g., Palikur *amepi-yo-* 'thief (woman)', *amepi-ye* 'thief (man)', Tariana *nu-phe-ri* 'my elder brother', *nu-phe-ru* 'my elder sister'. No genders are distinguished in the plural. The markers go back to proto-Arawak third person singular suffixes and prefixes: feminine *(r)u*, masculine *(r)i*. Some languages also have complicated systems of classifiers – these characterize the noun in terms of its shape, size, and function (Aikhenvald, 1999a). For instance, Tariana and Baniwa of Içana have more than 40 classifiers which appear on numerals, adjectives, verbs, and in possessive constructions. Palikur has more than a dozen classifiers which have different semantics and form depending on whether they are used on numerals, verbs, or on adpositions (Aikhenvald and Green, 1998). Pronominal genders have been lost from some languages, e.g., Terêna, Amuesha, Chamicuro, Pareci, Waurá (south of the Amazon), and Bahwana (north of the Amazon).

All Arawak languages distinguish singular and plural. Plural is only obligatory with human nouns. Plural markers are *-na/-ni* 'animate/human plural', *-pe* 'inanimate/animate non-human plural'. Dual

number is atypical. In Resígaro, markers of dual were borrowed from the neighboring Bora-Witoto languages.

Throughout the Arawak language family, nouns divide into those which must have a possessor (inalienably possessed) and those which do not have tohave a possessor (alienably possessed). Inalienably possessed nouns are body parts, kinship terms, and a few others, e.g., 'house' and 'name'. Inalienably possessed nouns have an 'unpossessed' form marked with a reflex of the suffix *-<i or *-hV, e.g., Pareci no-tiho 'my face', tiho-ti '(someone's) face'; Baniwa nu-hwida 'my head', i-hwida-ɫi (INDEFINITE-head-NON.POSSESSED) 'someone's head'. Alienably possessed nouns take one of the suffixes *-ne/ni, *-te, *-re, *-i/-e (Payne, 1991: 378), or *-na when possessed, e.g., Baniwa nu-<inu-ni (1sg-dog-possessive) 'my dog'.

The overwhelming majority of Arawak languages have a negative prefix ma- and its positive counterpart, prefix ka-, e.g., Piro ka-yhi (ATTRIBUTIVE-tooth) 'having teeth', ma-yhi (NEGATIVE-tooth) 'toothless'; Bare ka-witi-w (ATTRIBUTIVE-eye-FEMININE) 'a woman with good eyes', ma-witi-w 'a woman with bad eyes; a blind woman'.

The common Arawak lexicon (cf. Payne, 1991) consists mostly of nouns. There are quite a few body parts, fauna, flora, and artifacts. Only a few verbs can be reconstructed, e.g., *kau 'arrive', *pˆ(da) 'sweep', *po 'give', *(i)ya 'cry', *kama 'be sick, die'; *itha 'drink'. Most languages have just the numbers 'one' (proto-Arawak *pa-; also meaning 'someone, another') and 'two' (proto-Arawak *(a)pi and *yama). A preliminary reconstruction is in Payne (1991). An up-to-date overview of the family is in Aikhenvald (1999a, 2001), and an overview of the proto-language is in Aikhenvald (2002).

See also: Belize: Language Situation; Bolivia: Language Situation; Brazil: Language Situation; Classifiers and Noun Classes: Semantics; Colombia: Language Situation; Evidentiality in Grammar; Gender, Grammatical; Guatemala: Language Situation; Honduras: Language Situation; Nicaragua: Language Situation; Peru: Language Situation; Venezuela: Language Situation.

Bibliography

Aikhenvald A Y (1999a). 'The Arawak language family.' In Dixon R M W & Aikhenvald A Y (eds.). 65–105.

Aikhenvald A Y (1999b). 'Areal diffusion and language contact in the Içana-Vaupés basin, North West Amazonia.' In Dixon R M W & Aikhenvald A Y (eds.). 385–415.

Aikhenvald A Y (2001). 'Areal diffusion, genetic inheritance and problems of subgrouping: a North Arawak case study.' In Aikhenvald A Y & Dixon R M W (eds.). Areal diffusion and genetic inheritance: problems in comparative linguistics. Oxford: Oxford University Press. 167–194.

Aikhenvald A Y (2002). Language contact in Amazonia. Oxford: Oxford University Press.

Aikhenvald A Y & Green D (1998). 'Palikur and the typology of classifiers.' Anthropological Linguistics 40, 429–480.

Dixon R M W & Aikhenvald A Y (1999). 'Introduction.' In Dixon R M W & Aikhenvald A Y (eds.). 1–21.

Dixon R M W & Aikhenvald A Y (eds.) (1999). The Amazonian languages. Cambridge: Cambridge University Press.

Hoff B (1994). 'Island Carib, an Arawakan language which incorporated a lexical register of Cariban origin, used to address men.' In Bakker P & Mous M (eds.). Mixed languages. 15 case studies in language intertwining. Amsterdam: IFOTT. 161–168.

Payne D L (1991). 'A classification of Maipuran (Arawakan) languages based on shared lexical retentions.' In Derbyshire D C & Pullum G K (eds.). Handbook of Amazonian languages, vol. 3. Berlin: Mouton de Gruyter. 355–499.

Taylor D M (1977). Languages of the West Indies. Baltimore: John Hopkins University Press.

Tovar A & De Tovar C L (1984). Catálogo de las lenguas de América del Sur. Madrid: Editorial Gredos.

Architecture of Grammar

R Kempson, King's College London, London, UK

A grammar of a language is by definition a formal model of the properties of language that are intrinsic to the words of the language and the way they can be combined. There are then a number of choices to make:

1. Should the explanation be defined solely to reflect structural properties of language, or should those properties in some sense be correlated with the way language is used in speaking or hearing?

This is the issue of language competence and its separation from language performance.

There is another criterion which a grammar might be required to meet, which might seem like just another way of putting the same question, but is taken as requiring a different answer: