

## Introduction

W. Leo Wetzels (Paris III, Sorbonne Nouvelle/CNRS, LPP and Vrije Universiteit Amsterdam)

The articles in this volume originate from two sources. The manuscripts by Sérgio Meira & Spike Gildea, Francesc Queixalós, Masayoshi Shibatani & Khaled Awadh Bin Makhashen, and Willem Adelaar were selected by the editor. All others were written by researchers engaged in the Endangered Languages Programme financed by the Netherlands Organization for Scientific Research. The papers are organised according to the geographical distribution of the languages discussed, resulting in the following division, respectively: South America (Colombia, Suriname, Brazil and the Andes), Africa (Ethiopia, Ghana), and Asia (Yemen, East-Timor).

### *South America*

The language called Wánsöjöt Yedókjet, more often referred to as Puinave, is spoken by some 5000 individuals. These people are divided into two communities, one of which lives in the region of the Inírida River in Colombia, and the other, a smaller group, resides on the shores of the Venezuelan Orinoco. Wánsöjöt has sometimes been classified as belonging to the Makú-Puinave linguistic family, together with the Colombian languages Nikak, Kakua, and the eastern (mainly Brazilian) Maku languages. However, there is a lack of strong evidence for a genetic relationship between Wánsöjöt and any of these languages. Therefore, it must be considered a linguistic isolate for the time being. Although the Wánsöjöt represent a relatively numerous group, the number of speakers who master the traditional language is rapidly decreasing, mainly because over half of the population now live in urban areas and many children and adolescents opt for Spanish. In the article “Morfología y función de las construcciones nominalizadas en wánsöjöt (puinave)”, Jesús Mario Girón discusses the question of nominalization in Wánsöjöt, which has nominalized structures that function either as predicates or as participants. Deverbal nominalizations create agentive or objective nouns. In addition, action nominals and locative nominalizations may be

derived from verbs, which serve as oblique arguments. Nominalization is not limited to the lexicon; syntactic constituents may also be nominalized. Extended incomplete clauses and clauses that suffix the resultative *-pɣn* or the nominal past *-hin* function as subordinate and coordinate clauses. Furthermore, the conditional construction is formed by the addition of the locative case marker *-u* to the initial clause (protasis) of the conditional, whereas the past nominal *hin*, suffixed to one of two clauses in a sequence, creates adversative complex sentences. The Wánsöjöt data show a close relation between nominalization and predicates of relative clauses and other complement constituents.

The Nìkak and the Kakua once formed a cultural unit in the Vaupés region, Colombia, from which part of the population split off and migrated to the North. The 200 or so members of the Kakua people, also known as Bará-Maku, still live in southeast Colombia between the Querari and Papuri rivers, both tributaries of the Vaupés. It is the northern group that is known as the Nìkak, a nomadic people who were only officially contacted in 1988. Their number is estimated at roughly 550 individuals, all of whom live in the interfluvial region of the middle Guaviare and the upper Inírida. Soon after official contact, the Nìkak suffered a series of drastic changes in their way of life, most notably the reduction of about 40% of their population, their entrance onto the market as unskilled workers hired by the farmers who surround their land, and, since 2002, the massive displacement of various groups owing to the presence of armed forces in their territory. For these reasons, the survival of the Nìkak language and culture is not at all certain. In her paper “El nombre en nìkak”, Dany Mahecha Rubio discusses the phonological, morphological, and syntactic properties of nouns, nominalizations and nominal classification in Nìkak and compares nominalization and nominal classification with similar constructions in other languages of the region (Wánsöjöt, Yujup, Hup, Nadeb, and Dâw). In particular, she presents the few derivational processes the language has by which nouns are derived from verbs or nouns, and offers detailed observations on the structure of simple and complex nouns (both of which lack number and gender), the various types of compounds which may have nouns and verbs as their structuring elements, determiners, and case marking. Lastly, Mahecha Rubio provides arguments to distinguish two types of nominal classification.

The forty or so Cariban languages are spoken over an area that extends from the mouth of the Amazon River to the Colombian Andes and from Maracaibo, Venezuela to Central Brazil. Current estimates for speakers that are fluent in one of the Cariban languages range from 40,000 to 100,000. Given these comparatively impressive populations, the Cariban languages, together with Tupi and Arawak, represent one of the larger linguistic families of lowland South America. While most of the Tupi languages are spoken to the south of the Amazon River, the Cariban languages are mainly found to the north of it. The paper “Property Concepts in the Cariban Family: Adjectives, Adverbs, and/or Nouns?” by Sérgio Meira and Spike Gildea contributes to the typological issue of establishing what the universal lexical categories are, giving particular focus to the status of adjectives as a universal lexical category, as recently asserted by Dixon. Most modern descriptions of the Cariban languages have claimed that this language group does not have a separate category of adjectives, but that property concepts are divided between nouns and adverbs. Dixon reanalyzes data from reference grammars of three Cariban languages, claiming that the entire adverb category is better labeled “adjective”, and that the subset of property concept nouns is better considered a separate syntactic category of “Adjective2”. Meira and Gildea first demonstrate that Dixon’s specific analysis is unconvincing, but then that a closer look at the Cariban data yields a syntactic distinction between two subsets of the adverb class, one of which contains exclusively adjectival concepts and could, therefore, be considered a class of adjectives. They continue to suggest, however, that debating whether to call this a ‘subclass of adverbs’ or an ‘independent lexical category of adjectives’ is not an empirical, but rather a theoretical question (and therefore unanswerable on objective criteria alone). They instead investigate the diachronic origins of the unusual distribution of adjective concepts in the pan-Cariban lexical categories, arguing that this leads to a better understanding of the actual situation of “adjectives” and “adverbs” than simple binary classifications.

In her paper “Truth and Knowledge Markers in Wayana (Cariban), Suriname”, Eithne Carlin addresses the functionality of a set of grammatical markers in the Cariban language Wayana, spoken by some 1200 people spread out over an area that overlaps with borders between Suriname, French Guiana, and Brazil. The data presented are taken from the Surinamese Wayana, who live in the village of Píleoimë, along the Tapanahoni River. The Wayana harbour a negative attitude towards their

traditional language, directly attributable to the marginalized position of the indigenous communities vis-à-vis the national context. The Wayana in Suriname are rapidly learning Sranantongo, the lingua franca of the country, which is, however, still inadequate for improving their social standing especially in the state capital of Paramaribo, since Sranantongo, like Wayana itself, is a stigmatized language in formal settings. The grammar of Wayana contains a set of markers that includes a facsimile (similative) marker, several assertive or emphatic markers, a marker with the meaning ‘through and through’ or ‘truly’, a frustrative marker, nominal past suffixes, and evidential marking. Carlin concentrates on the non-verbal categories that express epistemological ideas of realities and truths. There is a pervasive regularity in Wayana and other Cariban languages of encoding states of being or not being, intimately connected with the Wayana world view. As Carlin shows, all the relevant markers, with the exception of the frustrative, express some temporal aspect of permanence vs. non-permanence, revealing an underlying temporal dimension ranging from simultaneity, progressivity, temporariness, to non-fluctuating permanence. Moreover, when nominal categories are marked by the frustrative, insights into cultural norms may be gleaned, but when it appears on verbs, it falls into the ‘knowledge’ domain.

Until fifty years ago, the nomadic and semi-nomadic Sikuni lived in small groups in the tropical savannas of eastern Colombia, between the Orinoco, Meta, Manacacias, and Vichada rivers. Beginning in the 1950s, some groups migrated toward the jungle areas of the Eastern Middle Orinoco in Venezuela and the Guaviare river in Colombia. Currently, the approximately 30,000 Sikuni populate both sides of the Middle Orinoco, with 70% of the population being in Colombia and the remainder in Venezuela. The Sikuni language is the largest of the small Guahibo family, which also contains the languages Cuiba (2,500 speakers), Guayabero (1,000 speakers), and Hitnü (250 speakers). It has been observed crosslinguistically that languages use terms related to the human body in a way that often transcends the naming of physical entities, properties, states or events. Sometimes they go as far as grammaticalizing the expressions in which these terms are embedded. The Sikuni language offers a nice illustration of this fact, as is shown by Francesc Queixalós in his paper “La posture du corps dans la classification et la localisation: l’exemple du sikuni”. There are four verbs describing the positions of the body that, on the one hand, construct a genuine categorization of the lexicon of nouns akin to more common nominal classifica-

tion systems, whereas, on the other hand, these verbs function to construct predications of localization and identification. Moreover, in their occurrences as auxiliaries, they go beyond the notion of space, expressing aspect and even modality meanings.

In his paper “Inverse markers in Andean languages: A comparative view”, Willem Adelaar compares the use of inverse markers in the verbal morphology of three unrelated Andean languages: Quechua, Puquina and Mapuche. The languages of the Quechuan linguistic family are spoken by some 8 million people in Peru and also in south-western and central Bolivia, the Andean region of Colombia and Ecuador, north-western Argentina, and northern Chile. Taken together, the Quechuan languages are the most widely spoken indigenous languages of the Americas. Puquina, a linguistic isolate, is now extinct. It was spoken in South Peru, on the border with Chili (Arequipa, Moquegua, Tacna). The Mapuche live in Central and Southern Chile and Southern Argentina. Of the more than one million Mapuche, who make up about 4% of the Chilean population, only some 200,000 may be considered fluent speakers. In Argentina, the Mapuche population is estimated at 60,000. As Adelaar argues, the data from Quechua, Puquina and Mapuche show a comparable development in their personal reference marking systems, in which inverse markers play a crucial role. Inverse markers allow languages like these, which display a limited set of personal reference endings (e.g. with subject markers only, or with an incomplete set of endings encoding both an actor and a patient in a transitive relation), to expand their inventory without having recourse to object markers specified for grammatical person. Instead, the absence of fully specified object markers is compensated by conveying the role of patient to what is normally a subject or agent marker. Inverse markers are used to indicate such a change of roles.

### *Africa*

Dime, Zargulla, and Sheko, spoken in Ethiopia, belong to the Omotic linguistic family, a primary branch of Afro-Asiatic, which also includes Berber, Chadic, Cushitic, Egyptian and Semitic. Like many other Omotic languages, Dime (with approximately 5,400 speakers) and Zargulla (with around 7,800 speakers) are endangered. Besides the low number of speakers, there is a strong socio-political pressure put on the minority ethnic groups of south-western Ethiopia in general, which contributes to the endangerment of their languages and cultural traditions. The Zargulla live to the west of Lake Chamo. Their lifestyle is similar to that of other

groups in the area, with the economy based on a balance between agriculture and herding. The Dime are sedentary and horticulturalists.

Two mutually intelligible dialects have been identified within the Dime society, Us'a and Gerfa. The Us'a dialect is the main source of data for Mulugeta Seyoum's paper "Negation in Dime". In Dime, negation is expressed through the use of the negative morpheme **-káy**. This morpheme also occurs following a copular verb. The negative nominal clause is headed by the obligatory negative copular verb **yi-** and, again, the negative marker **-káy**. In negative constructions, the aspectual and tense distinctions used in the affirmative are neutralized. Within the Omotic linguistic family, the neutralization of tense and aspect in the negative appears to be a special feature of Dime. On the other hand, it is a general feature of all the Omotic languages to use a special variant **-koy** of the negative morpheme for the negative imperative and optative.

In her paper "The morphosyntax of negation in Zargulla", Azeb Amha discusses the morphemes of the language that mark clausal negation; five of which are formally related and they are used in main clauses. The distribution of the five negative markers is determined by the predicate type (i.e. auxiliary, simple lexical verb or a complex predicate), tense-aspect and mood. Negation is marked by affixes attached to the verb in the non-past tense of declarative and interrogative clauses and in dependent clauses. In past declarative and interrogative clauses and in the imperative/optative mood however, it is expressed by independent negative verbs. There is partial similarity between the independent negative verbs and negative affixes, suggesting a (historical) link between the two. Amha attempts to account for the formal correspondences among these morphemes by identifying the direction of the historical derivation.

Yet another Omotic language reported on in this volume is Sheko, spoken by approximately 40,000 people, who live in the forested hills between Mizan Teferi and Tepi and on the Guraferda plateau in Southwest Ethiopia. Together with languages such as Dizi and Nayi, Sheko belongs to the so-called Dizoid group of languages, a subgroup of the North Omotic branch. In her paper "Possession in Sheko", Anne-Christie Hellenthal discusses the three constructions that the language possesses to express a possessor-possessum relation. Each of these constructions is of interest from a typological or historical-comparative point of view. Firstly, predicative possession in Sheko is similar to existential predication, in accordance with existing typological knowledge. However, the existential construction in Sheko permits only the possessum as a sub-

ject, contrary to what is found in other languages, for which predicative possession divides into two types, one of which has the possessor as a topic and the other the possessum. Secondly, attributive possession, in particular possessive noun phrases headed by **baab** 'father' or **bé** 'mother', show a peculiarity in gender marking, using *bé* irrespective of gender in definite phrases, even though masculine is the default gender in Sheko. Hellenthal also discusses possessor ascension in relation to body part nouns, taking into consideration inalienable possession in the related language Dizi. The author shows that the two constructions involved have different semantics, i.e. possessor ascension emphasizes the part (the possessum), while possessive noun phrases emphasize the whole (the possessor). Sheko and the other Dizoid languages remain atypical in marking inalienable possession by a case marker employing possessor ascension.

Logba and Tafi belong to the southernmost cluster of the Ghana-Togo Mountain (GTM) languages, a subgroup of the Niger Congo family. Logba and Tafi are two of fourteen different languages spoken in the hilly regions of the Ghana-Togo frontier, in the Volta Region of Ghana. Logba, spoken by roughly 7500 people, and Tafi, with approximately 4,400 speakers, are threatened by Ewe, the surrounding regional language to which both adults and children are shifting and (Ghanaian) English, the official language of Ghana.

In his paper "Noun Class System and Agreement Patterns in Logba (Ikpana)", Kofi Dorvlo studies the intricate noun class system of the language, where 'noun class' is understood as a set of nouns that can be characterized by 'a single set of morphological concords'. In noun class languages of the Niger Congo family in general, nouns have a particular prefix in the singular, while a different prefix is used in the plural. These languages usually also show a system of morphological concord between a noun and the verb. The GTM languages are reported by most researchers to have noun class systems and Logba is no exception as it shares the general noun class features of the Niger-Congo languages. Dorvlo argues that Logba noun classes correspond to semantic groupings, which contain subgroups with a somewhat looser semantic definition. There is a limited agreement within the head noun of the NP and a group of morphemes comprising a subset of the cardinal numerals (when used as modifiers), the demonstrative, and the interrogative. Moreover, there is general agreement between the subject NP and the verb, which is con-

trolled by the argument in subject position: the verbal agreement prefix is selected in function of the noun class that functions as the subject.

Just like Logba, Tafi has an active noun class system in which each noun belongs to a particular class identified by a prefix, as is shown by Mercy Lampsey Bobuafor in her paper “Noun Classes in Tafi: A Preliminary Analysis”. As in Logba, the head noun determines the type of affix that is used for agreement with the verb. The author argues that ten different noun classes must be distinguished for Tafi, based on the distribution of the class prefixes, the subject-verb agreement markers that are used to cross-reference the subject arguments on the verbs, and the shape of their pronominal forms. As in Logba, some modifiers show concord with the head noun while other word classes, like the adjectives and the ordinals, do not show such agreement. Some modifiers can be substantivized by adding class prefixes and the pronominal forms of the various noun classes to them and thus they function as nouns. The way borrowed nouns are integrated into the noun class system was also examined. Loans may also be allocated to a given class based on phonological and/or semantic criteria, or are assigned to the default word class.

#### *Asia*

Soqotri is a pre-Islamic Modern South Arabian language. With some 50,000 speakers, it is the native language of the Soqotrans, spoken on the islands of Soqotra, Abd-el-Kuri, and Samhah off the southern coast of Yemen. The language is regarded as endangered, since Arabic has become the official language on the island and because the youth overwhelmingly exhibit fluency in Arabic and an imperfect ability in their traditional language. The paper “Nominalization in Soqotri, a South Arabian language of Yemen” by Masayoshi Shibatani and Khaled Awadh Bin Makhshen argues for the nominalization analysis for several of the so-called headless NP constructions including the ones known as headless relative clauses. Contrary to the popular deletion analysis for expressions analogous to *John's* (as in *This book is John's*), *(We must help) the poor*, and *(I am reading) what John recommended to me*, arguments are advanced for the analysis of treating these as grammatical nominalizations, which in their basic function play a referring function. Soqotri, along with a number of other languages such as Chinese, Japanese and Lahu, makes use of the same nominalization morphology for all these constructions. An important distinction, however, is drawn between event nominalizations, on the one hand, and argument nominalization



and what is called genitive nominalization. Soqotri presents a clear case where different types of argument nominalization function as a modifier element of the relative clause construction, suggesting that relative clauses are nominalization-based rather than involving full subordinate sentences or clauses as in the standard analysis.

Fataluku is a non-Austronesian language spoken in the Lautém district in the eastern part of the Republic of East Timor. It genetically belongs to the non-Austronesian Timor-Alor-Pantar branch of the Trans-New-Guinea phylum. The Fataluku speech community is estimated to be comprised of around 30,000 people. Although the constitution of East Timor acknowledges that all the languages spoken on the territory of the republic must be protected, the promotion of Tetum as the language of communication in education and public life and its status as the co-official language (alongside Portuguese) has caused a shift from Fataluku to Tetum. It is for this reason combined with the redistribution of power (and prestige) in the newly created state, that the transmission of Fataluku to the younger generation is seriously declining. In his contribution to this volume “On Derivational Processes in Fataluku, a Non-Austronesian Language in East-Timor”, Aone van Engelenhoven discusses the (very few) productive derivational morphological processes of the language. The main nominalizing device of the language is a nominalizing morpheme /-(n)u/, which may be suffixed to the same roots as the verbalizing morpheme /-(n)e/. These mechanisms, which are mutually exclusive, give rise to formations like *tupur-u* ‘woman’ versus *tupur-e* ‘be a woman, feminine’. Formations of this type raise the question as to whether the roots to which these suffixes attach are pre-categorical. One of the few productive morphological processes of Fataluku takes state verbs as its input and creates outputs that function as modifiers (adverbs) of to the following verb. The same process of deverbal adverbialization creates serial verb constructions in which the derived adverb fills the object slot of the bivalent verb. The noun referring to the instrument, product, or theme participant of the scene evoked by the verb is encoded as the object of ‘take’ in the directly preceding clause.

Within the remit of the general theoretical debate, the contribution of theory-directed data is crucially important to improve data-oriented theory formation. For any kind of theoretical linguistics, the discovery of new linguistic patterns is a matter of great importance, since the empirical validity of theoretical models is rightly put to the test continually by

confrontation with the balance as it is actually found to exist between both the striking similarities between languages and also the differences between them. If considered from this perspective, the usefulness of the study of languages that are at the edge of disappearance is self-evident. It is the shared hope of all the contributors to this book that the topics discussed here serve to emphasize the importance of data from endangered or dying languages to further develop the field of linguistics.