The Mekong-Mamberamo mystery

David Gil[1], Sihan Chen[2], and Antonio Benítez-Burraco[*3]

[*]Corresponding Author: abenitez8@us.es

[1]Department of Linguistic and Cultural Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambrisge, MA, USA

[3]Department of Spanish, Linguistics & Theory of Literature, University of Seville, Spain

This paper suggests that linguistic areas, or sprachbunds, may constitute relics of earlier stages in the evolution of language. Here we focus on the *Mekong-Mamberamo* linguistic area, extending from Mainland Southeast Asia though the Indonesian archipelago and into western New Guinea. The first part of this paper surveys evidence that Mekong-Mamberamo languages exhibit a distinctive grammatical profile associated with greater simplicity in both morphology and syntax. The second part of this paper examines potential explanations for the simple grammatical profile associated with the Mekong-Mamberamo area, and concludes that the most likely of these is that it constitutes an evolutionary relic from an earlier stage in the evolution of language.

1. Introduction

In studying the phylogeny of language, one common method is to try to identify features of contemporary languages that might constitute models for an earlier stage in the evolution of human language. Such evolutionary relics may potentially be present in a number of different domains. First, they may be found embedded in the architecture of particular subsystems of grammar. Thus, Progovac (2015) argues that small clauses, and various other defective clause types, identifiable as part of the more elaborate syntactic structures of languages such as English and Serbo-Croatian, may be viewed as fossils from an earlier evolutionary stage of language. Secondly, certain language types may be viewed more holistically as representative of earlier stages in the evolution of language. Thus, for example, Gil (2017) claims that some contemporary languages, such as Riau Indonesian, come close to instantiating an idealized IMA language — Isolating (lacking internal word structure), Monocategorial (lacking distinct parts of speech), and Associational (lacking construction-specific rules of semantic

compositionality) — potentially representing a model for an earlier stage in the evolution of language. Similarly, further down the evolutionary line, Benítez-Burraco and Progovac (2020) suggest that contemporary languages spoken by esoteric, inward-oriented societies are characterized by a more complex morphology alongside a simpler syntax, and that such languages may thus also be considered to represent an earlier stage in the phylogeny of language.

This paper proposes a third domain in which such an evolutionary relic may be observed, namely the linguistic area, or sprachbund. To date, sprachbunds have been mostly used to reconstruct deep stages of language change (e.g. Bickel and Nichols, 2006), but not to infer the types of languages putatively spoken in our remote past. Here we focus on the *Mekong-Mamberamo* linguistic area, first introduced in Gil (2015). This sprachbund consists roughly of mainland Southeast Asia, the Indonesian archipelago, and western parts of the island of New Guinea. Its name derives from the two major rivers located at its two extremities, the Mekong to the north, and the Mamberamo to the east. We survey a body of evidence showing that, compared to a worldwide baseline, Mekong-Mamberamo languages are typically, and sometimes by a substantial margin, associated with greater simplicity with regard to their morphological and syntactic structures. The simple grammatical profile of Mekong-Mamberamo languages poses a mystery: How and why did this profile come into being? This paper suggests that the most likely explanation is that the simple grammatical profile of Mekong-Mamberamo languages is an areally-defined relic representing an earlier stage in the evolution of language.

2. The simple grammatical profile of Mekong-Mamberamo languages

As proposed in Gil (2015), the Mekong-Mamberamo linguistic area is motivated by 17 typical linguistic features. Of these, 7 can be considered as entailing greater grammatical simplicity:

- low differentiation of adnominal attributive constructions
- weakly developed grammatical voice
- isolating word structure
- short words
- low grammatical-morpheme density
- optional thematic-role flagging
- optional Tense-Aspect-Mood (TAM) marking

The Mekong-Mamberamo linguistic area is illustrated with reference to the last of these 7 features, TAM marking, in the following map:

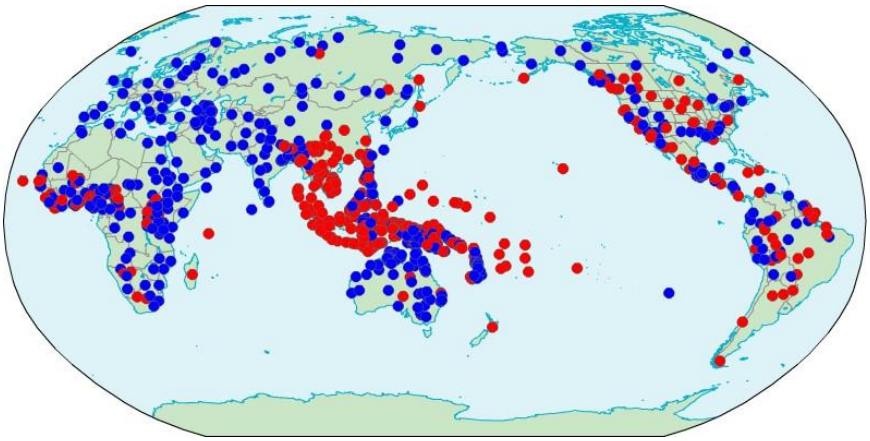

Figure 1. Optional (●) and obligatory (●) Tense-Aspect-Mood marking.

Eyeballing the above map shows clearly that the Mekong-Mamberamo linguistic area, right in the middle of the map, is the only one in which TAM marking is consistently optional. In other parts of the world, such as sub-Saharan Africa, Oceania and the Americas, optional and obligatory TAM-marking languages are interspersed, while in central and western Eurasia and north Africa, obligatory TAM-marking languages are the rule. These geographical patterns are analyzed in more detail in Gil (2021).

Further independent support for the simple grammatical profile of the languages of the Mekong-Mamberamo area derives from a wider study about potential trade-offs between morphology and syntax in the world languages (Benítez-Burraco et al., 2024). In this study, based on the 144 grammatical features listed in WALS, we selected 44 features pertaining to morphological complexity and 39 features pertaining to syntactic complexity. Complexity here is understood in purely descriptive terms: if a grammatical value requires more description than some other vale of the same feature, it is considered as more complex (e.g. Li and Vitányi, 2008; Sinnemäki, 2011). Furthermore, since assigning a grammatical feature to either morphology or syntax can be tricky, and may depend on background theoretical assumptions about the nature of grammar (and even language), we followed the simplest criterion possible: if a grammatical feature pertains to word structure, it was considered as a morphological feature, whereas if it pertains to relationships between words, it was considered as a

syntactic feature. To assign each language morphological and syntactic complexity scores, we averaged the normalized values across features pertaining to morphology and syntax respectively. However, due to the limited data availability in WALS, languages vary dramatically in terms of feature coverage. In this study, we considered only the 461 languages in the WALS database for which sufficient data is available.

Figure 2 shows the results of our analysis, plotting the syntactic complexity (S) scores of the 461 languages against the morphological complexity (M) scores, and with the Mekong-Mamberamo languages highlighted in red.

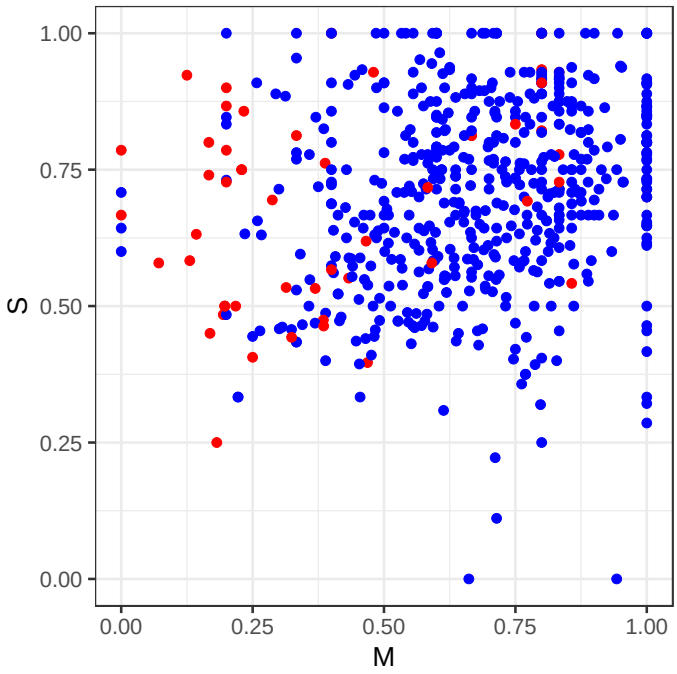

Figure 2. Complexity of Mekong-Mamberamo (●) and other (●) languages. Here the Mekong-Mamberamo area is taken to consist of China south of 30N, all the countries of Mainland Southeast Asia, the Indonesian archipelago (including Timor Este) but not the Philippines, plus New Guinea and associated islands west of 135E.

To evaluate our hypothesis that, following Gil (2015), languages within this area exhibit a simpler grammatical profile, we adopted a Monte-Carlo-based approach, sampling one language from the Mekong-Mamberamo area and another from outside the area and then comparing their morphological complexity and syntactic complexity scores. The 95% confidence interval we obtained was [0.50, 0.67],

significantly higher than 0.25 (the threshold for a by-chance relationship). Overall, the results supported our hypothesis.

3. Towards a resolution of the mystery

The Mekong-Mamberamo linguistic area thus presents us with a mystery: How could a large geographical region have come to be associated with a grammatical profile so systematically different (and simpler) from that found in other parts of the world? In Gil (2015:412-4), a number of speculative answers are put forward, appealing to factors such as language contact and the relatively recent presence of other hominin species; however, none of these answers enjoys clear cut support. In lieu of such support, our default hypothesis is that the Mekong-Mamberamo linguistic area is the way that it is because it always was like that. In other words: the simple grammatical profile characteristic of Mekong-Mamberamo languages is a relic from an earlier stage in the evolution of language in which languages exhibited less complex morphologies and syntaxes.

As is well known, the diachronic accretion of complexity is a ubiquitous process, observable throughout the world even in relatively shallow time depths. In view of this, one may indeed wonder how likely it is that a geographical region as large as the Mekong-Mamberamo might have been spared such pervasive diachronic processes of complexification for such a long time, allowing for the preservation of an earlier stage in the evolution of language itself. Consideration of contact between closely related dialects or languages reveals some of the mechanisms that might have contributed to the perseverance, over time, of the simpler Mekong-Mamberamo grammatical profile (Gil 2020:190-1). For example, in dialects of Kerinci spoken in central Sumatra, most words occur in one of two competing forms, absolute and oblique; however, in the emerging Kerinci koiné, the absolute/oblique alternation is in the process of disappearing, under influence from surrounding Mekong-Mamberamo languages. Processes such as these suggest that the simpler grammatical profile associated with Mekong-Mamberamo languages may be self-perpetuating, providing a second pole of stability around which languages may cluster and persevere.

An apparent challenge to the archaic nature of the simple Mekong-Mamberamo profile derives from the diachronic study of the Austronesian languages occupying a large central swathe of the area. It is commonly accepted that the original grammatical profile of Proto-Austronesian was the more complex one that is currently observable in contemporary Austronesian languages of Taiwan and the Philippines, and that the simpler grammars of many of the other Austronesian languages of the Indonesian archipelago are due to contact-induced simplification that took place some 3,500-4,000 years ago, when Austronesian languages spread south into the archipelago. According to this view, their simpler

grammatical profiles would actually be an innovation dating back just a few thousand years at the most. However, in Gil (2020) it is argued that the Austronesian languages of the Indonesian archipelago are most appropriately viewed as exhibiting dual heritage, reflecting the coming together of two distinct linguistic lineages: while the vocabulary is largely Austronesian, much of the grammar represents a direct inheritance from the non-Austronesian languages that were already present in the region. In other words, the simpler grammatical profiles associated with today's Austronesian languages of the Indonesian archipelago predate the arrival of Austronesian languages in the region; their presence in the region was a continuous one, dating back as far as we can see.

Clearly, at an earlier point in human evolution, languages were simpler than they are today. Accordingly, the plausibility of the hypothesis that the Mekong-Mamberamo linguistic profile is an evolutionary relic depends on how far back in human pre-history one must go until all the world's languages exhibited the simpler grammatical profiles of today's Mekong-Mamberamo languages. Consideration of the worldwide geographical distribution of grammatical complexity suggests that this might have been at a relatively recent stage, post-dating the spread of modern humans out of Africa. As argued in Gil (2009), a simple IMA language is all that was needed to facilitate collective tasks such as sailing a boat to an island over the horizon. Thus, humans could have spread out all over the world, speaking languages associated with a simple grammatical profile resembling that of contemporary Mekong-Mamberamo languages. Later, complexification would have occurred, arising independently in several locations, and then spreading until it encompassed most of the world — with the exception of the Mekong-Mamberamo area. In the same vein, Benítez-Burraco and Progovac have hypothesized that humans spoke simpler languages perhaps as late as 50.000 years ago, at which time languages began to complexify under the effects of our increased prosocial behavior. All in all, in accordance with such scenarios, the Mekong-Mamberamo linguistic area may have been left behind as a relic of an earlier stage in the evolution of language. It remains to be determined which specific factors, seemingly extralinguistic by nature (social, cultural, or even environmental), contributed to preserve these relic features in this part of the world.

References

Bickel, B., & Johanna (2006) "Oceania, the Pacific Rim, and the theory of linguistic areas". In *Annual Meeting of the Berkeley Linguistics Society*, 32(2), 3-15.

Benítez-Burraco, Antonio and Ljiljana Progovac (2020) "A four-stage model for language evolution under the effects of human self-domestication", *Language and Communication 73.1–17*

Chen, Sihan, David Gil, Sergey Gaponov, Jana Reifegeste, Tessa Yuditha, Tatiana Tatarinova, Ljiljana Progovac and Antonio Benitez-Burraco (2023) "Linguistic and Memory Correlates of Societal Variation, A Quantitative Analysis", PsyArXiv Preprints, 10.31234/osf.io/bnz2s.

Gil, David (2009) "How Much Grammar Does It Take to Sail a Boat?", in G. Sampson, D. Gil and P. Trudgill eds., *Language Complexity as an Evolving Variable*, Oxford University Press, Oxford, 19-33.

Gil, David (2015) "The Mekong-Mamberamo Linguistic Area", in N.J. Enfield and B. Comrie eds., *Languages of Mainland Southeast Asia, The State of the Art*, Pacific Linguistics, DeGruyter Mouton, Berlin, 266-355.

Gil, David (2017) "Isolating-Monocategorial-Associational Language", in H. Cohen and C. Lefebvre eds, Handbook of Categorization in Cognitive Science, Second Edition, Elsevier, Amsterdam, 471-510.

Gil, David (2020) "Dual Heritage: The Story of Riau Indonesian", in D. Gil and A. Schapper eds., *Austronesian Undressed: How and Why Languages Become Isolating*, John Benjamins, Amsterdam, 119-212.

Gil, David (2021) "Tense-Aspect-Mood Marking, Language Family Size, and the Evolution of Predication", in A. Benítez-Burraco and L. Progovac eds., Prehistorical Languages, *Philosophical Transactions B*.

Progovac, Ljiljana (2015). *Evolutionary Syntax*. Oxford: Oxford University Press.