

How Much Grammar Does It Take to Sail a Boat?

David Gil

Max Planck Institute for Evolutionary Anthropology, Leipzig

(to appear in G. Sampson, D. Gil and P. Trudgill eds, *Language Complexity as an Evolving Variable*, Oxford University Press, Oxford)

first draft, October 2007

1. COMPLEXITY OF LANGUAGE, COMPLEXITY OF CIVILIZATION

Human languages are of much greater complexity than the communicative systems of great apes, dolphins, bees and other animals. Similarly, human culture, technology and civilization are also immensely more complicated than anything observed in other species, such as, for example, the ways in which chimpanzees fashion tools to crack nuts or fish for termites. Clearly, these two facts are related: comparing humans to other species leads inexorably to the conclusion that linguistic complexity is *correlated* with complexity in other, non-linguistic domains. But what exactly is the nature of this correlation?

A widespread assumption is that linguistic complexity is necessary in order to support complexity in other domains. This accords with a functional approach towards the evolution of language, whereby greater linguistic complexity enables humans to accomplish more tasks, and in doing so confers an evolutionary advantage. It also forms the basis for archaeological investigations into the evolution of language, in which the existence of material artifacts from a certain era is interpreted as evidence that humans at that time were also endowed with the cognitive and linguistic abilities necessary for the production of such artifacts (Lee and DeVore 1968, Clark 1977, Roebrooks 2001, de la Torre 2004 and many others). Similarly, the discovery of the remains of an apparently new species of hominin on the Indonesian island of Flores raises the possibility that such hominins were capable of constructing and sailing boats, which in turn may suggest that *Homo Floresiensis* was endowed with whatever linguistic abilities are necessary for the collective planning and execution of such a complex set of activities (see Morwood and Cogill-Koez 2007 for recent critical discussion).

But what, exactly, are the necessary linguistic abilities: How much grammar does it really take to build a boat and sail it to a distant island? Or more generally: How much grammar does it take to do all of the things that, amongst all living species, only humans are capable of doing, such as, for example, worshiping god, waging war, and working in offices; inventing, manufacturing and using sophisticated hi-tech tools; and engaging in multifarious commercial, scientific and artistic activities.

This paper argues that the amount of grammar that is needed in order to support the vast majority of daily human activities is substantially less than is often supposed to be the case, in fact less than that exhibited by any contemporary human language, and far less than that exhibited by most such languages. In other words, much of the observable complexity of contemporary human grammar has no obvious function pertaining to the development and maintenance of modern human civilization. More specifically, it is argued that the level of grammatical complexity necessary for contemporary culture, technology and civilization is no greater than that of *Isolating-Monocategorical-Associational* (or *IMA*) Language.

2. ISOLATING-MONOCATEGORIAL-ASSOCIATIONAL LANGUAGE

Isolating-Monocategorial-Associational Language, introduced in Gil (2005a), is an idealized language prototype with the following three properties:

- (1) (a) *Morphologically Isolating*
No word-internal morphological structure;
- (b) *Syntactically Monocategorial*
No distinct syntactic categories;
- (c) *Semantically Associational*
No distinct construction-specific rules of semantic interpretation
(instead, compositional semantics relies exclusively on the Association Operator, defined in (2) below).

As argued in Gil (2005a, 2006), IMA Language characterizes an early stage in the phylogenetic evolution of human language, and also an early stage in the ontogenetic development of contemporary child language. In addition it can be observed in artificial languages such as that of pictograms. However, no known contemporary human language instantiates IMA Language; all such languages are endowed with additional kinds of structures beyond those sanctioned by the definition in (1) above.

The defining properties of IMA Language represent the limiting points of maximal simplicity within each of three distinct domains, morphology, syntax and semantics. Hence, for each domain, one may imagine languages approaching these end points along a scale of decreasing complexity. Accordingly, a language is increasingly isolating as it has less and less morphological structure, increasingly monocategorial as its syntactic categories decrease in number and importance, and increasingly associational as its construction-specific rules of semantic interpretation become fewer and less distinct. Alongside *Pure IMA Language*, as in (1) above, one may thus entertain the possibility of a range of *Relative IMA Languages*, approaching Pure IMA Language to various degrees within each of the above three domains.

The first defining property, *morphologically isolating*, is the one that is most familiar, since it forms the basis for a typology that has been the focus of considerable attention in the linguistic literature. As is well known, isolating languages such as Vietnamese have considerably less word-internal morphological structure than synthetic languages such as Russian, which in turn have considerably less morphology than polysynthetic languages such as Mohawk. However, no natural language is purely isolating; all known isolating languages still have *some* morphology — affixation, compounding, or other kinds of processes such as reduplication, stem alternation, and so forth.

The second defining property, *syntactically monocategorial*, pertains to a domain within which the presence of cross-linguistic variation has only recently, and still only partially, been recognized. In the past, syntactic categories have generally been presumed to be universal; however, in recent years an increasing body of literature has begun to examine the ways in which the inventories of syntactic categories may vary across languages. One major area of focus has been the case of languages purportedly lacking in various parts-of-speech distinctions, such as languages with no adjectives, or languages in which nouns and verbs cannot be distinguished. However, to the best of my knowledge, no language has ever actually been proposed to be purely monocategorial. In

particular, most or all descriptions of languages still involve, at the very least, a distinction between one or more open syntactic categories containing "content words" and one or more closed syntactic categories containing various "grammatical" or "functional items"; see, for example, Gil (2000).

The third defining property, *semantically associational*, relates to the domain of compositional semantics, the way in which the meanings of complex expressions are derived from the meanings of their constituent parts. This property makes reference to the *Association Operator*, defined as follows:

(2) *The Association Operator A:*

Given a set of n meanings $M^1 \dots M^n$, the Association Operator A derives a meaning $A (M^1 \dots M^n)$ read as 'entity associated with M^1 and ... and M^n '.

Setting n equal to 1 results in the *Monadic Association Operator*, manifest, for example, in constructions containing a genitive, possessive or associative marker. Allowing n to equal 2 or more results in the *Polyadic Association Operator*, in accordance with which, whenever two or more expressions group together to form a larger expression, the meaning of the combined expression is associated with, or has to do with, the meanings of each of the individual expressions. The Polyadic Association Operator is responsible for the fact that in a two-word expression in which, say, one of the words contains a stem meaning 'chicken' while the other contains a stem meaning 'eat', the meaning of the expression as a whole must be related to 'chicken' and 'eat', for example 'The chicken is eating', 'The chickens that were eaten', 'The reason chickens eat'; it can never mean 'Beavers build dams'. The Polyadic Association Operator is thus a universal default mechanism for semantic interpretation, albeit one that is in most cases overridden and narrowed down substantially by the application of additional construction-specific rules. A purely associational language would be one in which there were no such further construction-specific rules of semantic interpretation, and in which, therefore, the compositional semantics were effected exclusively by the Polyadic Association Operator. It is almost certainly the case that no natural language is purely associational; however, as argued in Gil (2007, to appear), some languages make relatively less use of such additional semantic rules, and may thus be characterized as more highly associational.

3. RIAU INDONESIAN AS A RELATIVE IMA LANGUAGE

No naturally occurring contemporary human language completely satisfies the definition of IMA Language. However, whereas many languages, such as English, Hebrew, Dani and Pirahã, go way beyond the confines of IMA Language, exhibiting much greater levels of complexity, others approach the IMA prototype to various degrees, thereby warranting characterization as Relative IMA Languages. One example of a Relative IMA Language is Riau Indonesian, as described in Gil (1994, 2000, 2001, 2002a,b, 2004a,b, 2005a,b and elsewhere).

In Riau Indonesian, basic sentence structure is in fact purely IMA. Consider the following simple sentence:

- (3) Ayam makan
chicken eat
A (CHICKEN, EAT)

The above sentence consists entirely of two "content words", and is devoid of any additional grammatical markers. The *isolating* character of the language is instantiated by the fact that each of the two words is monomorphemic. The *monocategorial* nature of the language is reflected by the fact that the two words, although referring to a thing and an activity respectively, exhibit identical grammatical behaviour; rather than belonging to distinct parts of speech, such as noun and verb, they are thus members of the same syntactic category, namely sentence, and therefore the sentence as a whole is a simple juxtaposition, or coordination, of two sentences. The *associational* character of the language can be seen in the wide range of available interpretations: the first word, *ayam*, is underspecified for number and definiteness; the second word, *makan*, is indeterminate with respect to tense and aspect; and the sentence as a whole is underspecified with regard to thematic roles, with *ayam* being able to bear agent, patient, or any other role in relation to *makan*, and in addition also indeterminate with respect to ontological categories, with possible interpretations belonging to categories such as activity, thing, reason, place, time and others. However, although the sentence can be understood in very many different ways, such as 'The chicken is eating', 'The chickens that were eaten', 'The reason chickens eat' and so forth, it is not multiply ambiguous; rather, it is extremely vague, with a single very general interpretation which may be represented, as in (3) above, with the Polyadic Association Operator, A (CHICKEN, EAT), to be read as 'entity associated with chicken and eating', or, more idiomatically, 'something to do with chicken and eating'. Sentence (3) above is a typical sentence in Riau Indonesian; it is not "telegraphic" or otherwise stylistically marked in any way. Longer and more complex sentences can be constructed which, like (3), instantiate pure IMA structure.

Nevertheless, Riau Indonesian contains a number of features which take it beyond the bounds of a pure IMA Language. That Riau Indonesian is not purely isolating is evidenced by the presence of a handful of bona fide affixes, plus various other morphological processes, such as compounding, reduplication and truncation. None of these processes, however, are inflectional, and none of them are obligatory in any particular grammatical environment. That Riau Indonesian is not purely monocategorial is due to the fact that in addition to the single open syntactic category sentence, it also contains a single closed syntactic category containing a few dozen semantically-heterogeneous words whose grammatical behaviour sets them apart from words

belonging to the category of sentence. However, although most of members of this second, closed syntactic category are what are generally considered to be grammatical function words, none of these items are obligatory in any specific grammatical construction. Finally, that Riau Indonesian is not purely associational is clear from the presence of additional rules of compositional semantics that make reference to specific lexical items or to particular syntactic configurations, such as, for example, word order. Still, the effect of such rules is much more restricted than is the case in many other languages. Thus, Riau Indonesian is most appropriately characterized as a Relative IMA Language.

4. IMA LANGUAGE IS ALL THAT'S NEEDED TO SAIL A BOAT

Riau Indonesian is a colloquial language variety used in informal situations as a vehicle of interethnic and increasingly also intraethnic communication by a population of about 5 million in the province of Riau in east-central Sumatra. Riau Indonesian is but one of a wide range of colloquial varieties of Malay/Indonesian, spoken throughout Indonesia and neighboring Malaysia, Brunei and Singapore by a total population of over two hundred million people. Although differing from each other to the point of mutual unintelligibility, a majority of these colloquial varieties resemble Riau Indonesian in their basic grammatical structures, and accordingly share the characterization as Relative IMA Languages. However, in addition to these basilectal language varieties, there are also versions of Standard Indonesian and Malay, typically acquired in mid-childhood or later as a second or subsequent language variety for use in formal situations; as might be expected, these more acrolectal varieties of Malay/Indonesian are less IMA than their colloquial counterparts.

As Relative IMA Languages, Riau Indonesian and other colloquial varieties of Malay/Indonesian make it possible to address the question: How much grammar does it take to sail a boat? By peeling off the extra layers of non-IMA complexity and homing in on the IMA core, one may examine the expressive power of pure IMA Language, and see exactly what levels of culture, technology and civilization it can support. In order to do this, we shall take a look at some fragments of pure IMA Language culled from naturalistic corpora in a few varieties of colloquial Malay/Indonesian. In such corpora, most utterances contain at least some additional structure beyond what is purely IMA: an affix, a word belonging to the closed class of grammatical function words, or a construction-specific rule of semantic interpretation. Nevertheless, it is also possible to find stretches of text in which, by probabilistically-governed accident, no such additional structure is present, and in which, therefore, exhibit pure IMA structure. Following are some such examples of pure IMA text:¹

¹ The interlinear glosses make use of the following abbreviations: CONTR contrastive; DEM demonstrative; DIST distal; IPFV imperfective; NEG negative; PL plural; PROX proximal; PST past; Q question; SG singular; 1 first person; 2 second person; 3 third person.

- (4) "Ck, emang lu libur besok?" tadi kan Erto.
 tsk really 2 holiday tomorrow PST:PROX Q Erto
 'Then Erto asked, right "Tsk, do you really have a day off tomorrow?"'
- (5) Tadi ogut pas telefon, Eto lagi mandi.
 PST:PROX 1:SG precise telephone Eto more wash
 'Right when I phoned, Eto was taking a shower.'
- (6) Kenapa kamu ngga pengen ikut?
 why 2 NEG want follow
 'Why don't you want to come along?'
- (7) Semester tujuh sampe skarang semester delapan ni.
 semester seven arrive now semester eight DEM:PROX
 Semester delapan tinggal sampe kemarin baru bayar SPP.
 semester eight remain arrive yesterday new pay tuition.fee
 Baru bayar SPP selesai tra tau tong masuk kulia
 new pay tuition.fee finish NEG know 1:PL enter university.class
 kapan ni.
 when DEM:PROX
 'From the seventh semester until now, the eighth semester. Through the eighth semester until yesterday, we only just paid the tuition fee. But even though we just paid the tuition fee, we don't know when we're going to start classes again.'
- (8) Baru de dapa sepak langsung kaki pata.
 new 3:SG get kick direct leg break
 Jadi semester enam tu de cuti.
 become semester six DEM:DIST 3:SG holiday
 'Then he was kicked and broke his foot. So he was out for semester six.'
- (9) Pas SMA tu sa bilang "mama kam beli sa gitar
 precise senior.high.school DEM:DIST 1:SG say mother 2:PL buy 1:SG guitar
 baru dulu". Dong beli sa gitar baru suda mo.
 new before 3:PL buy 1:SG guitar new IMPF want
 'When I went into senior high school I said "Mum, you should buy me a new guitar". So they bought me a new guitar.'

- (10) Kan dia pengen orang luar nikah samo dio, samantaro.
 Q 3 want person out marry together 3 temporary
 Orang tuo nyo kan ndak buliah,
 person old 3 Q NEG can
 dio pingin anak nyo nikah samo anak Datung.
 3 want child 3 marry together child Datung
 'She wanted an outside person to marry her, temporarily, right. But her parents
 wouldn't allow it, right, they wanted their child to marry the son of Datung.'
- (11) Karam tu. Ntah ndak tau wak ntah
 collapse DEM:DIST NEG.know NEG know 1 NEG.know
 a karam rumah tu pas urang duduak ateh rumah
 what collapse house DEM:DIST precise person sit up house
 'It collapsed. I don't know how the house could collapse while people were sitting
 in it.'
- (12) Udah tu, hari ujan. Hari ujan masuk lah,
 IMPF DEM:DIST day rain day rain enter CONTR
 ado gua dakek kampuang tu, masuk lah gua tu nyo.
 exist cave near village DEM:DIST enter CONTR cave DEM:DIST 3
 Masuk dalam gua tu kan, paruik lah lapa ko.
 enter inside cave DEM:DIST Q abdomen IPFV hungry DEM:PROX
 'After that, it rained. As it was raining, he went in, there was a cave near the
 village, and he went into the cave. Having gone into the cave, he was hungry.'
- (13) Korsi kami korsi kayu,
 chair 1 chair wood
 Korsi miko korsi buloh;
 chair 2:PL chair bamboo
 Orang kami orang Dayak,
 person 1 person Dayak
 Orang miko orang Batak.
 person 2:PL person Batak
 'Our chairs are wooden chairs / Your chairs are bamboo chairs / Us people are
 Dayak people / You people are Batak people.'

Examples (4) - (6) above are from Jakarta Indonesian, the colloquial variety of Indonesian spoken as the primary everyday language by over 20 million residents of the capital city of Indonesia and its surrounding metropolitan area, and increasingly adopted as a stylistically trendy register by middle and upper class inhabitants in the provinces. Examples (7) - (9) are from Papuan Malay, used by possibly 2 million people as the lingua franca in the western half of the island of New Guinea under Indonesian control. Examples (10) - (12) are from Minangkabau, spoken by an estimated 6-7 million members of the eponymous ethnicity in the province of West Sumatra as well as in

migrant communities elsewhere. (Minangkabau is usually considered to be a "different language", however it is very closely related to Malay/Indonesian, and not much more different from many varieties of Malay/Indonesian than such varieties are from each other; for ease of exposition, we shall thus subsume it under the heading of colloquial Malay/Indonesian.) And example (13) is from Siak Malay, a rural dialect spoken by a few hundred thousand people in the Siak river basin in Riau province in east-central Sumatra.²

The dialects represented in the above examples span a variety of sociolinguistic types. Whereas Jakarta Indonesian and Papuan Malay are the possible products of radical restructuring such as that characteristic of creole languages, and are now spoken by ethnically mixed populations, Minangkabau and Siak Malay are more appropriately viewed as direct descendants of proto-Malayic, and now constitute emblematic vehicles of the respective ethnic identities. Speakers of these different varieties range from westernized and upwardly mobile office workers in high-rise buildings in Jakarta, through shop-keepers and rice-farmers across the archipelago, all the way to New Guinea highlanders in penis gourds and grass skirts. Together, these four varieties of colloquial Malay/Indonesian provide a representative sample of the unity in diversity inherent in what is by all accounts a major world language.

The above examples provide an indication of the expressive power of pure IMA Language, showing how it matches up to other, non-IMA Languages, by capturing notions that in other languages make recourse to specialized grammatical constructions. Demonstratives occur in examples (7), (8), (9), (11) and (12); numerals in (7); and negatives in (6), (7), (10) and (11) — all completely within the confines of IMA Language, without any of the structurally specialized constructions used to express the corresponding notions in many other languages. Tense and aspect are expressed in pure IMA structure in several of the above examples — proximate past *tadi* in (4) and (5), progressive *lagi* in (5), sequential *baru* in (7) and (8), and perfectives *suda* in (9) and *udah* and *lah* in (12). And words with first, second or third person reference but not belonging to a grammatically dedicated category of pronoun can be found in each and every one of the above examples.

Other, more relational notions are also expressed within the confines of pure IMA Language. Various semantic types of attribution are all expressed by simple juxtaposition of head and modifier: *gitar baru* 'new guitar' in (9), *anak nyo* 'their child' and *anak Datung* 'Datung's child' in (10), and eight other instances in (13). The formation of a content question is exemplified in (6); the shift in perspective characteristic of a passive construction is illustrated in (8); while some of the ways in

² Jakarta Indonesian has been the subject of a number of recent investigations, including Wouk (1989, 1999), Cole, Hermon and Tjung (2006) and Sneddon (2006). Papuan Malay has featured less prominently in the literature, one recent study being that of Donohue and Sawaki (2007); it is also discussed briefly in Gil (2002b), where it is referred to as Irian Indonesian. Minangkabau is described in a reference grammar by Moussay (1981). And Siak Malay has been discussed in the context of its contact relationship with Riau Indonesian in Gil (2001, 2002a, 2004a). The other authors mentioned above do not necessarily subscribe to the present characterization of these language varieties as Relative IMA Languages.

which coreference can be maintained are shown in (9), (10) and (12) — all remaining within the limits of pure IMA Language.

Complex meanings which in other languages call for various kinds of embedded clauses may also be expressed within pure IMA Language. Among these are temporal clauses in (5), (9) and (11); complements of mental acts in (10); embedded polar questions in (4); embedded content questions in (7) and (11); reported speech in (4) and (9); and tail-head linkage, in which the final sequence of an utterance is repeated, in backgrounded form, at the beginning of the next utterance, in (7) and (12). In fact, example (11), with its temporal clause occurring within the embedded question, shows how such meanings can be nested, thereby underscoring the recursive potential of pure IMA Language.

Thus, as suggested by the above examples, pure IMA Language is endowed with substantial expressive power. In fact, comparing the above pure IMA fragments to the totality of the Relative IMA Language varieties from which they are taken suggests that getting rid of the non-IMA accoutrements — as per the above exercise — has no systematic effect on expressive power in any semantic domain. The affixes and grammatical markers that take these language varieties beyond the bounds of pure IMA Language form a semantically heterogeneous set, a functional hodge-podge sprinkled like confetti over their fundamentally IMA architecture. A few examples should make this point clear. First, consider the well-known voice markers *di-* and *N-*, which, as prefixes or proclitics (depending on the dialect and on the phonological properties of the host word) go beyond pure isolating and monocategorial structure. However, unlike prototypical passive and active voice markers, *di-* and *N-* have no effect on the grammatical structure of the clause; they are thus optional semantic embellishments (Gil 2002b). Next, consider the marker *yang*, often described as a relative-clause marker. Although colloquial varieties of Indonesian do not have bona fide relative clauses, the grammatical behaviour of *yang* — it can only occur in front of another expression which functions as its host — places it outside the single open syntactic category and therefore beyond the limits of monocategoriality. However, its use is always optional; for example, 'the chickens that were eaten' can be rendered as either *ayam yang makan*, or, simply, as in (3) above, *ayam makan*. Finally, consider forms with meanings corresponding to English prepositions, such as *dengan* 'with'. Like *yang*, it can only occur in front of a host expression and is therefore not a member of the single open syntactic category of sentence. However, any expression containing *dengan* can be paraphrased with an alternative expression not containing *dengan* and falling within the scope of pure IMA Language. Often, such paraphrases will contain the word *sama*, whose range of meanings wholly contains that of *dengan*, but which differs from *dengan* in being a member of the single open syntactic category of sentence (Gil 2004b). Examples such as these demonstrate that the various non-IMA grammatical markers do not make much of a difference when it comes to assessing the overall expressive power of colloquial Malay/Indonesian varieties. In principle, anything that can be said in such languages can be paraphrased within the confines of pure IMA Language.

This being the case, pure IMA fragments such as those in (4) - (13) paint a reasonably accurate picture of the functionality of IMA Language, and the amount of culture, technology and civilization that IMA Language can support. In a nutshell: IMA Language is enough to run a country of some two hundred million people, and, by

extension, most contemporary human activity throughout the world. Most of the examples in (4) - (13) are conversational, but this merely reflects the fact that conversation is the most common human linguistic activity, and the one that our corpora accordingly focus upon. Still, as suggested by these examples, the restriction to IMA Language does not impose any constraints on the range of things that can be talked about, or on what can actually be said about those things. As the above examples show, IMA Language is enough to talk about school, sports, love and life. And it is sufficient to support most daily activities throughout one of the world's largest countries, from the giant metropolis that is Jakarta to the most far flung of island provinces. Moreover, lest this come across as too utilitarian a view of language, example (13) — a *pantun*, the most widespread genre of oral poetry in Malay/Indonesian — demonstrates that IMA Language can also provide the raw material for verbal art.

Admittedly, there are contexts, mostly of a more formal and official nature, where the colloquial language is inappropriate, and instead the somewhat more complex standard language is used. This raises the possibility that there may exist some domains for which the Relative IMA character of colloquial Malay/Indonesian is functionally inadequate. However, in many cases, at least, the use of the standard language is motivated not by any functional gain in expressive power, but rather by social conventions. For example, the president addressing the nation on television could easily get his message across in colloquial Jakarta Indonesian, but to do so would result in loss of face, thereby endangering his elevated standing. One important domain in which the standard language is typically preferred over colloquial varieties is that of writing. However, it is striking that although most Indonesians nowadays can read and write, Indonesia remains a functionally illiterate society: people prefer to communicate orally rather than in writing. A ubiquitous example is provided by public transportation, where almost every vehicle — bus, minibus, boat or whatever — has somebody hanging on to the outside and calling out the destination to prospective passengers; even if there is also a written sign conveying the same information, people still listen rather than look. Thus, even if there do exist some contexts where the greater grammatical complexity of the standard language really plays a necessary communicative role, something that is not at all obvious, such contexts are marginal and few in number, paling into insignificance alongside the so many more domains for which IMA Language proves sufficient to do the job.

What colloquial Malay/Indonesian shows us, then, is that IMA Language is all that it takes to sail a boat. This means that, if indeed Homo Floresiensis sailed across a major body of water to reach the island of Flores, the most that we can infer from this with regard to his linguistic abilities is that he had IMA Language. More generally, what this suggests is that no amount of non-linguistic archeological evidence will ever be able to prove the presence, at some earlier stage of human evolution, of grammatical entities of greater-than-IMA complexity: prefixes and suffixes, nouns and verbs, not to mention complementizers, relative clauses, government, movement, functional categories, antecedents binding empty positions, and all the other things that so delight the souls of linguists.

5. WHY IS GRAMMAR SO COMPLEX?

If indeed IMA Language is all it takes to sail a boat and to run a large country, why is it

then that no languages are pure IMA Languages, and most languages are not even Relative IMA Languages, instead exhibiting substantially greater amounts of grammatical complexity? One cannot but wonder what all this complexity is for. This paper has only been able to provide a negative answer, by identifying one albeit enormous thing that this complexity is not for, namely, the maintenance of contemporary human civilization: IMA Language is enough for all that.

As noted at the outset, comparing humans to other species suggests that grammatical complexity is in fact positively correlated with complexity in other non-linguistic domains. However, more fine-toothed observations within the human species reveal a more ambivalent picture. Admittedly, within certain specific contexts, it is possible to identify what appear to be significant correlations between grammatical complexity and complexity in other domains, as for example in colloquial Malay/Indonesian, where the development of a non-IMA and hence more complex coordinator may be shown to be related to the introduction of mobile telephony and text messaging (Gil 2004b). However, in other contexts, the correlation seems to go in the opposite direction, as in the well-known case of language simplification being associated with the greater sociopolitical complexity of contact situations (McWhorter 2005, Trudgill this volume, and others). Thus, across the diversity of contemporary human languages and cultures, grammatical complexity just does not seem to correlate systematically with complexity in other, non-linguistic domains. In the words of Sapir (1921:219): "Both simple and complex types of language of an indefinite number of varieties may be spoken at any desired level of cultural advance. When it comes to linguistic form, Plato walks with the Macedonian swine-herd, Confucius with the head-hunting savage of Assam."

These facts cast doubt on a central tenet of most functionalist approaches to language, in accordance with which grammatical complexity is there to enable us to communicate the messages we need to get across. In spite of overwhelming evidence showing that diachronic change can be functionally motivated, the fact remains that language is hugely dysfunctional. Just think of all the things that it would be wonderful to be able to say but for which no language comes remotely near to providing the necessary expressive tools. For example, it would be very useful to be able to describe the face of a strange person in such a way that the hearer would be able to pick out that person in a crowd or a police line-up. But language is completely helpless for this task, as evidenced by the various stratagems police have developed, involving skilled artists or, more recently, graphic computer programmes, to elicit identifying facial information from witnesses — in this case a picture actually being worth much more than the proverbial thousand words. Yet paradoxically, alongside all the things we'd like to say but can't, language also continually forces us to say things that we don't want to say; this happens whenever an obligatorily-marked grammatical category leads us to specify something we would rather leave unspecified. English, famously, forces third person singular human pronouns to be either masculine or feminine; but in many contexts we either don't know the person's gender or actually wish to leave it unspecified, and we are thus faced with the irritating choice between a stylistically awkward circumlocution such as *they*, *he or she*, or (in writing) *s/he*, or, alternatively, the non-politically-correct generic *he* or its overly-politically-correct counterpart *she*. Examples such as this can be multiplied at will. What these facts suggest, then, is that grammatical structure with its

concomitant complexity is not a straightforward tool for the communication of pre-existing messages, but that rather, to a large degree, our grammars actually define the messages that we end up communicating to one another.

Instead of wondering what grammatical complexity is for, one should ask how and why grammars have evolved to their current levels of complexity. Clearly, many factors are involved, some common to all languages, underlying the development of grammatical complexity in general, others specific to individual languages, resulting in the observed cross-linguistic variation with respect to grammatical complexity. Among the many different factors involved, a particularly important role is played by diachrony. Contemporary grammatical complexity is the result of thousands of years of historical change, with its familiar processes of grammaticalization, lexicalization, syntacticization and the like. Rather than having evolved in order to enable us to survive, sail boats, and do all the other things that modern humans do, most contemporary grammatical complexity is more appropriately viewed as the outcome of natural processes of self-organization whose motivation is largely or entirely system-internal. In this respect, grammatical complexity may be no different from complexity in other domains, such as anthropology, economics, biology, chemistry and cosmology, which have been suggested to be governed by general laws of nature pertaining to the evolution of complex systems.

ACKNOWLEDGEMENTS

Most of the naturalistic data discussed in this paper were collected by the staff at the Max Planck Institute Jakarta Field Station and its Padang and Manokwari branches. For their administrative and professional assistance I am indebted to Uri Tadmor and Betty Litamahuputty (Jakarta), Yusrita Yanti (Padang), and Yusuf Sawaki (Manokwari); and for collecting and transcribing the data I am grateful to Sarah Chakrawati, Erni Farida Sri Ulina Ginting, Ferdinan Okki Kurniawan, Dalan Mehuli Perangin-angin and Tessa Yuditha (Jakarta), Santi Kurniati, Silvie Antisari Anhar and Yessy Prima Putri (Padang) and Fitri Djamaludin, Yehuda Ibi, Nando Saragih and Olivia Waren (Manokwari). For the *pantun* I am thankful to Kudin (Sungai Pakning).

REFERENCES

- Clark, Graham (1977) *World Prehistory in New Perspective*, Cambridge University Press, Cambridge.
- Cole, Peter, Gaby Hermon and Yassir Nasanius Tjung (2006) "Is There *Pasif Semu* in Jakarta Indonesian?", *Oceanic Linguistics* 45:64-90.
- Donohue, Mark and Yusuf Sawaki (2007) "Papuan Malay Pronominals: Forms and Functions", *Oceanic Linguistics* 46:253-276.
- Gil, David (1994) "The Structure of Riau Indonesian", *Nordic Journal of Linguistics* 17:179-200.
- Gil, David (2000) "Syntactic Categories, Cross-Linguistic Variation and Universal Grammar", in P. M. Vogel and B. Comrie eds., *Approaches to the Typology of Word Classes*, Empirical Approaches to Language Typology, Mouton, Berlin and New York, 173-216.
- Gil, David (2001) "Creoles, Complexity and Riau Indonesian", *Linguistic Typology* 5:325-371.

- Gil, David (2002a) "Ludlings in Malayic Languages: An Introduction", in Bambang Kaswanti Purwo ed., *PELBBA 15, Pertemuan Linguistik Pusat Kajian Bahasa dan Budaya Atma Jaya: Kelima Belas*, Unika Atma Jaya, Jakarta, 125-180.
- Gil, David (2002b) "The Prefixes *di-* and *N-* in Malay/Indonesian Dialects", in F. Wouk and M. Ross eds., *The History and Typology of Western Austronesian Voice Systems*, Pacific Linguistics, Canberra, 241-283.
- Gil, David (2004a) "Learning About Language from Your Handphone; *dan*, *and* and *&* in SMSs from the Siak River Basin", in Katharina Endriati Sukatmo ed., *Kolita 2, Konferensi Linguistik Tahunan Atma Jaya*, Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya, Jakarta, 57-61.
- Gil, David (2004b) "Riau Indonesian *sama*, Explorations in Macrofunctionality", in M. Haspelmath ed., *Coordinating Constructions*, Typological Studies in Language 58, John Benjamins, Amsterdam, 371-424.
- Gil, David (2005a) "Isolating-Monocategorial-Associational Language", in H. Cohen and C. Lefebvre eds., *Categorization in Cognitive Science*, Elsevier, Oxford, 347-379.
- Gil, David (2005b) "Word Order Without Syntactic Categories: How Riau Indonesian Does It", in A. Carnie, H. Harley and S.A. Dooley eds., *Verb First: On the Syntax of Verb-Initial Languages*, John Benjamins, Amsterdam, 243-263.
- Gil, David (2006) "Early Human Language Was Isolating-Monocategorial-Associational", in A. Cangelosi, A.D.M Smith and K. Smith eds., *The Evolution of Language, Proceedings of the 6th International Conference (EVOLANG6)*, World Scientific, Singapore, 91-98.
- Gil, David (2007) "Creoles, Complexity and Associational Semantics", in U. Ansaldo and S.J. Matthews eds., *Deconstructing Creole: New Horizons in Language Creation*, John Benjamins, Amsterdam, 67-108.
- Gil, David (to appear) "How Complex Are Isolating Languages?" in F. Karlsson, M. Miestamo and K. Sinnemäki eds., *Language Complexity: Typology, Contact, Change*, John Benjamins, Amsterdam.
- Lee, Richard B. and Irven DeVore (1968) *Man the Hunter*, Aldine, Chicago.
- McWhorter, John (2005) *Defining Creole*, Oxford University Press, Oxford and New York.
- Morwood, Michael J. and Dorothea Cogill-Koez (2007) "*Homo* on Flores: Some Early Implications for the Evolution of Language and Cognition", in A.C. Schalley and D. Khlentzos eds., *Mental States, Volume 1: Evolution, Function, Nature*, Studies in Language Companion Series 92, John Benjamins, Amsterdam.
- Moussay, Gerard (1981) *La langue Minangkabau*, Archipel, Paris.
- Roebrooks, Wil (2001) "Hominid Behavior and the Earliest Occupation of Europe: An Exploration", *Journal of Human Evolution*, 41:437-461.
- Sapir, Edward (1921) *Language*, Harcourt, Brace and World, London.
- Sneddon, James Niel (2006) *Colloquial Jakartan Indonesian*, Pacific Linguistics, Canberra.
- Torre, Ignacio de la (2004) "Omo Revisited: Evaluating the Technological Skills of Pliocene Hominids", *Current Anthropology* 45:439-465.
- Trudgill, Peter (this volume) "Sociolinguistic Typology and Complexification".
- Wouk, Fay (1989) *The Impact of Discourse on Grammar: Verb Morphology in Spoken Jakarta Indonesia*,. PhD Dissertation, UCLA, Los Angeles.

Wouk, Fay (1999) "Dialect Contact and Koineization in Jakarta, Indonesia", *Language Sciences* 21:61-86.