

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/272816392>

# Emergent Grammar

Article in Proceedings of the Annual Meeting of the Berkeley Linguistics Society · September 1987

DOI: 10.3765/bls.v13i0.1834

---

CITATIONS

329

---

READS

4,566

1 author:



Paul Hopper

Carnegie Mellon University

91 PUBLICATIONS 6,556 CITATIONS

SEE PROFILE

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



Anacrustic Coordination [View project](#)

## Emergent Grammar

Paul Hopper

'Fragments are the only forms I trust.'  
- Donald Barthelme, *See the Moon?*

'Essence is expressed by grammar.'  
- Ludwig Wittgenstein, *Philosophical Investigations* 371

### 1. Emergent Grammar<sup>1</sup>

As explorations in 'functional grammar' accumulate in volume and significance, it has become a standard tactic of supporters of sentence syntax to claim that the very study of discourse is an unreasonable agenda so long as any problems remain outstanding from the study of sentence level syntax. This claim sometimes takes the form of challenges to functional grammar to find discourse correlates of specific syntactic phenomena (stated always, of course, in sentence terms). For example, we find Jerry Morgan asking the question how such well-known phenomena as Ross' constraints on variables could conceivably be explained in discourse terms (Morgan 1981). Indeed, how could the 'extraposition' of the relative clause in such ordinary everyday sentences as:

'The woman died in 70,000 BC who invented the wheel,'

ever be accounted for if the structure of every sentence in the language had to have a functional explanation?

The same theme of the arbitrariness of the match between syntax and function is struck by other linguists who have confronted discourse linguistics from the perspective of sentence grammar; we can cite here Newmeyer's pronouncement of a so-called 'functionalist fallacy' (Newmeyer 1983), the position that the lamination between structure and function will come unglued with the slightest bubble of failure—that it takes only one syntactic fact which is not susceptible of a functional explanation to bring down the whole precarious palace of functionalism. Sadock (1984) has endorsed the notion of a Functionalist Fallacy.

There are, it seems to me, both superficial and more profound responses to this argument. The simple one is to note that since both sentence grammarians and discourse linguists agree that their work is not done, it is highly premature to speak of any 'syntactic facts' which are independent of function, just as in the absence of a complete theory of

functionalism it is premature to claim that all structures have functional counterparts. It might also be legitimately claimed that so-called syntactic facts which do not appear at first sight to have a functional explanation may indeed have one if they are studied seriously, that is, in real discourses. Thus in the particular example presented by Morgan, it could be argued that extraposition of relative clauses is indeed triggered by something in discourse, probably involving the relative salience of the main vs. the subordinate clause: extraposition always seems to mean that the discourse importance of the relative clause outranks that of the main clause; indeed it is the very absence of this skewing which perhaps accounts for something I always notice when I try this sentence out on people — that it is judged to be a very bizarre way to say what the sentence is apparently trying to say.

Critics of ‘radical pragmatics’, and ‘functional grammar’, assume that they and those they oppose share a common view of language, that there is a pairing of autonomous (i.e., decontextualized) grammatical forms with ‘functions’ (whatever they might be in the abstract), and that the only point of disagreement is whether these forms might be eventually derivable from ‘functions’ or whether the forms must be described independently of ‘functions’. I find a certain irony in such a use of the terms ‘function’ and ‘functionalism’, since the very restriction of the investigation to an artificially defined level of ‘sentences’ seems to me to be quintessentially anti-functional. Be that as it may, I am concerned in this paper with the more fundamental problem of the assumptions underlying the critique, especially the assumption of an abstract, mentally represented rule system which is somehow implemented when we speak.

It is an assumption which is very deeply entrenched in our field, and indeed is virtually an official dogma. Consider the following. A year or so ago, the President of the LSA, Victoria Fromkin, was asked by the editors of the *Chronicle of Higher Education* to submit a brief state-of-the-art report on linguistics, to be featured in a two-page spread of similar reports by representatives of other disciplines. Here is part of what Fromkin wrote:

‘In human speech production and comprehension, the speaker-hearer accesses not only the mentally represented language system, but also other cognitive systems and knowledge of the world.’ (Fromkin 1985: 13)

A whole world of unarticulated philosophical and other assumptions underlies this statement. It is a world whose traces are glimpsed through terminological windows such as ‘access’, ‘mental’, ‘representation’, ‘language system’, and ‘cognitive system’. And that is only the intellectual aspect of the statement, for we must not forget that it is simultaneously a political statement also, a public inscription by the President of the Linguistic Society of the boundaries and objectives of the field of linguistics. But I am concerned more with the basic scenario, the one which provides for a logically prior — perhaps eventually even biologically prior — linguistic system which is simultaneously present for all speakers and hearers, and which is a pre-requisite for the actual use of language. It is, in other words, the scenario that when we speak we refer to an abstract, mentally represented rule system, and that we in some sense ‘use’ already available abstract structures and schemata.

The assumption, in other words, is that 'grammar' (in the sense of the rules, constraints, and categories of the language attributed to the speaker) must be an object apart from the speaker and separated from the uses which the speaker may make of it. That kind of grammar is conventionally understood to consist of sets of rules which operate on fixed categories like nouns and verbs, specify the forms of additive categories like those of case, tense, transitivity, etc., and restrict the possible orders in which words can occur in a sentence. Discourse, the actual use of language, is held to be in some sense an 'implementation' of these structures, or the way in which the abstract mental system possessed in its entirety by the speaker is realized in particular utterances.

Discourse linguistics has itself not always been immune to this kind of thinking. Here, too, one frequently encounters the same assumption of a dualistic structure in discourse, the notion that structure pre-exists discourse and that discourse is mimetically related to a logically prior abstract organization, formulated this time in terms of paragraphs, episodes, events, and other such macro-units. The problems of sentence grammar are not really alleviated by treating discourses as units manifesting a consistent internal structure, in other words effectively as extra-long sentences. We are still plagued by the problem of the ill-fitness of fit between form and function. However consistently it can be predicted that a certain particle or aspectual form will function in a particular role in the discourse, it is rare that the reverse is the case—that a particular form is restricted to a single specifiable discourse role. To cut a very long story short, and thereby probably caricature the dilemma, some way out of the vicious circle of form-to-function-to-form is needed.

This is, then, roughly the context in which the term Emergent Grammar is being proposed. The term 'emergent' itself I take from an essay by the cultural anthropologist James Clifford, but I have transferred it from its original context of 'culture' to that of 'grammar'. Clifford remarks that 'Culture is temporal, emergent, and disputed' (Clifford 1986:19). I believe the same is true of grammar, which like speech itself must be viewed as a real-time, social phenomenon, and therefore is temporal; its structure is always deferred, always in a process but never arriving, and therefore emergent; and since I can only choose a tiny fraction of data to describe, any decision I make about limiting my field of inquiry (for example in regard to the selection of texts, or the privileging of the usage of a particular ethnic, class, age, or gender group) is very likely to be a political decision, to be against someone else's interests, and therefore disputed.

The notion of Emergent Grammar is meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process. Grammar is hence not to be understood as a pre-requisite for discourse, a prior possession attributable in identical form to both speaker and hearer. Its forms are not fixed templates but are negotiable in face-to-face interaction in ways that reflect the individual speakers' past experience of these forms, and their assessment of the present context, including especially their interlocutors, whose experiences and assessments may be quite different. Moreover, the term Emergent Grammar points to a grammar which is not abstractly formulated and abstractly represented, but always anchored in the specific concrete form of an utterance.

The notion of emergence is a pregnant one. It is not intended to be a standard sense of origins or genealogy, not a historical question of 'how' the grammar came to be the way it 'is', but instead it takes the adjective emergent seriously as a continual movement towards structure, a postponement or 'deferral' of structure, a view of structure as always provisional, always negotiable, and in fact as epiphenomenal, that is at least as much an effect as a cause.

The assumption of Emergent Grammar imposes on the linguist a rather radically different view of the data base for linguistics. Although isolated, made-up clauses and sentences will have their uses, and indeed are often indispensable short-cuts to the study of grammar, the sources of these forms will have to be understood in a different way from that of the fabula of abstract rules and native speaker intuitions which have become part of our dogma. The linguist's task is in fact to study the whole range of repetition in discourse, and in doing so to seek out those regularities which promise interest as incipient sub-systems.

Structure, then, in this view is not an overarching set of abstract principles, but more a question of a spreading of systematicity from individual words, phrases, and small sets. I will illustrate what I mean by considering the example of the English indefinite article *a/an*. If we consider the history of this form, we find that from Indo-European times a cognate form has meant the simple numeral 'one', singularity. This was still a common meaning of *án* in Old English. It is seen in such examples (from Bosworth and Toller 1898, sub *án*) as:

God geworhte ánnan mannan, Adam, of láme  
'God created one man, Adam, out of clay'

It is also commonly used to introduce a new participant into a discourse:

án man hæfde twegen suna  
A [certain] man had two sons'

Its use as a general indefinite article does not appear until later, so that in Old English *án* is not found in such contexts as:

Deodric wæs Cristen  
'Theodoric was a Christian'

To take just these three functions of the predecessor of *a/an* in Old English, we find in modern English not a uniform, over-all weakening of the meaning, but rather a situation in which the weakened meanings and the older stronger meanings exist side-by-side. Thus we find, among other uses, the indefinite sense of a non-specific, classifying article:

My husband and I went to a showroom to buy a new car. After we had  
test-driven one, the salesman asked us ... (from *Redbook*)

But the specific, new-mention sense is also found:

They introduced me to a young woman [whose name was Ethel].

My husband and I went to a showroom to pick up a new car we had ordered. After we had test-driven it, the salesman asked us ...

Although these senses — specific and non-specific — have usually been taken as exhaustively dividing up the domain of the indefinite article, in fact several other uses also exist, such as ‘one and the same’:

Birds of a feather flock together  
They are all of a kind

and even ‘one’:

A stitch in time saves nine  
A penny saved is a penny earned  
How much is that picture-frame? — A dollar.

—where, by the way, British English would require ‘one dollar’. It is significant that these meanings of ‘one’ and ‘the same’ are not replicable outside of the contexts — and in some cases the specific wording — of these formulas. Thus discourses like the following seem anomalous:

Linguists of a theory attend the same conferences.  
What was left of the woods after they built the parking-lot? — A tree.

Evidently the meanings represented by the English ‘indefinite article’ are not unified under one hyper-abstract function. Instead, an open ended set of small sub-systems has come into being, and the membership of new occurrences of forms with the indefinite article is not specifiable in advance, but is impromptu and negotiable. *Even participants in the conversation may not know whether a specific new mention or a non-specific indefinite is intended until this has been worked out in the verbal interaction.* Moreover, these subsystems are either innovating and spreading out from an earlier more restricted usage, or are contracting and being abandoned from an earlier wider use. We see this most clearly in the obviously traditioned, formulaic diction of proverbs like ‘birds of a feather’, where ‘a feather’ retains not only the older sense of ‘one and the same’, but also the singular noun ‘feather’ in the sense of plumage. The spread of the newer, indefinite-nonspecific function of *a/an* was described in Hopper and Martin 1987.

The point about the retention of archaisms in proverbial language has of course often been made. But it has less often been noted that proverbial language is only an extreme case of repetition in discourse, at the other end of which are the morphological and syntactic repetitions some of which are called grammar; this point is made cogently by Lambrecht 1984. In other words, real live discourse abounds in all sorts of repetitions which have nothing to do with grammar as this is usually understood: for instance, idioms, proverbs,

clichés, formulas, specialist phrases, transitions, openings, closures, favored clause types, and so on. There is no consistent level at which these regularities are storable. They are not necessarily 'sentences', or 'clauses', with recurrent internal structure, but they are often used holistically. Their boundaries may or may not coincide with the constituent boundaries of our grammatical descriptions: subject and predicate, noun phrase, prepositional phrase. Moreover, what is a formulaic expression in one context may not be in another; again, see Lambrecht 1980.

It has been noted before that to a very considerable extent everyday language is built up out of combinations of such prefabricated parts. Language is, in other words, to be viewed as a kind of pastiche, pasted together in an improvised way out of ready-made elements. Language is thus to be treated, in Wittgenstein's words, 'from outside' (cf. Wittgenstein 1958: para. 120) — not as governed by internalized mentally represented rules, but by pre-existent material with which discourses can be devised; cf. Staten 1984: 85-86, Smith 1978: 61-62 et pass. Evidently an entirely parallel way of viewing language is to be attributed to Jacques Derrida with his metaphor of language as 'graft': new speech acts are 'grafted onto' old ones and of course serve in turn as the stock onto which further new speech acts are crafted (cf. Culler 1982: 134-135). Becker's idea of 'prior texts' (e.g., Becker 1979: 244-245) is also crucial here: previous actual utterances form the basis of new utterances. Similar observations have been made by Bolinger, by Andrew Pawley, and others. It is this pre-patterned, pre-fabricated aspect of speech which accounts best for the characteristic of language for which no dualistic, double-tiered theory can provide an intuitively satisfying explanation: in natural discourse we compose and speak simultaneously (Smith 1978: 60). There is no room — no need — for mediation by mental structures. It is in this sense that, as Bolinger has pointed out (Bolinger 1976), speaking is more similar to remembering procedures and things than it is to following rules. It is a question of possessing a repertoire of strategies for building discourses and reaching into memory in order to improvise and assemble them. Grammar is now not to be seen as the only, or even the major, source of regularity, but instead grammar is what results when formulas are re-arranged, or dismantled and re-assembled, in different ways.

Looking at language this way involves a serious adjustment for the linguist, since we have developed the habit of seeing utterances in terms of a fixed framework of rules, and especially because we have been raised on the doctrine of the free generability of sentences, and the privileging of novelty over prior texts. Indeed, novelty is a prized virtue in our society altogether,<sup>2</sup> and we have many ways, some more subtle than others, of censuring perceived repetitions of others' behavior and an enormous vocabulary dealing with repetition (copying, imitation). Yet when one examines actual specimens of speech from the formulaic point of view the effect is a striking one, perhaps even a memorable one, in that it is then extremely difficult to revert to the old rule-governed syntactic view of discourse. Consider the following example from spoken English, just one of many examples from the Carterette and Jones corpus:

Well no the problem is and this is what the psychologist has mentioned  
to me. these kids wont wont show any hope like the see you take a

normal uh the average retarded child i mean the one who doesnt have any handicaps like blindness or deafness or something like that. he will improve a little bit. maybe a lot. it depends on how badly disturbed he is. but these people wont because theyre still going to no matter what happens theyre going to be living in a fantasy world. because theyre blind. and they have to imagine and they keep asking one question after the other and then nothing they say makes any sense and nothing is relevant to the situation. and it never will be because they well theres just such a sharp line of differentiation between the normal blind and then the emotionally disturbed blind. (Carterette and Jones 1974:422).

Even a cursory study of such passages reveals several different layers of regularity. The formulas are easily isolated. Note just a few of them:

the problem is  
has mentioned to me  
these kids  
you take  
a little bit  
maybe a lot  
it depends on  
no matter what happens  
theyre still going to  
living in a fantasy world  
one question after another  
nothing they say makes any sense  
relevant to the situation  
sharp line of differentiation  
emotionally disturbed

—the last with its institutional and authoritarian subtext. It would in fact be difficult or impossible to draw the line between a formulaic and a non-formulaic expression; Moreover, there are single words which could themselves be said to constitute formulas in this context, such as ‘disturbed’, ‘normal’. The stops and starts coincide with the boundaries of formulas, which are presented and modified or withdrawn or capitalized upon in an obvious interactional negotiation. Early in the paragraph, for example, the speaker clearly is about to say ‘the normal retarded child’, but some way into the phrase realizes that for the uninitiated in this context it clashes with another formula, ‘the normal child’, and launches into a second try, ‘the average retarded child’, which also — once said — appears incongruous (cf. ‘the average child’), and finally is forced to abandon the search for an appropriate formula and move into a more specific level of discourse in which the properties encapsulated in the adjective ‘normal’ are made explicit:



these kids wont wont show any hope like the see you take a normal uh  
the average retarded child i mean the one who doesnt have any  
handicaps like blindness or deafness or something like that.

It might be suggested that in this particular passage a sort of second hand 'health care professional' jargon is manifest, in which mannerisms peculiar to a particular set of experts intrude. (We might note, for example, the pervasive 'will/won't' in place of the present tense, right out of the H.C.P.'s manual!) Yet it would be difficult to find a passage about which some analogous remark might not be made. The point is that all discourse is in some sense specialist discourse, molded to the speaker's personality (i.e., personal history), the situation, and the topic. It is precisely the point about Emergent Grammar that such 'heteroglossic' (Bakhtin 1981:281) aspects of language necessarily become integral parts of the linguistic description, and are not set aside as a separate agenda irrelevant to the linguistic code and its structure.

Some of these phrases are incongruous when considered from a structural-grammar perspective. Consider the phrase beginning 'You take a...', which would have to be analyzed as a subject-verb construction. Its actual function in the monologue is quite different from what one might predict of such a phrase on structural grounds. It is not a report ascribing an action of taking to a second person subject. It has in fact only one function, expressed holistically: to present a new hypothetical case into the discourse context. But this function cannot be readily integrated into a homogeneous grammatical system whose postulates obtain only at the level of isolated sentences and which starts with the perspective of a solitary ideal speaker.

The systematicity which linguists have come to expect in language exists, of course, but in a more complex way. The linguistic system is now not to be seen as something complete and homogeneous, in which 'exceptional' phenomena must be set aside as inconvenient irregularities, but as a *growing together* of disparate forms. This convergence takes place through lateral associations of real utterances. Similarities spread outwards from individual formulas, in ways that are motivated by a variety of factors, such as:

- (i) phonological similarity (rhyme, assonance): he's likely to —> he's liable to
- (ii) contextual similarity: I persuaded him to —> I convinced him to,

and other kinds of resonance.<sup>3</sup> They do not, however, merge into the kind of uniform grammar which would lead one to posit a uniform mental representation to subtend them.

## 2. Preferred Clauses

2.1 What I've been saying up to now has had the purpose of recontextualizing the notion of grammar—not to abolish it, but rather to suspend it with a view to isolating those regularities in discourse which we will agree to call emergent grammatical regularities. But as we have seen, the doctrine of Emergent Grammar assigns an entirely different status to grammar from what might be called A Priori Grammar:

(1) Regularity in discourse is of many different kinds, and is, since there is continually movement between one kind and another, moreover dynamic, not static in nature. Consequently no principled line can be drawn between the emergent regularities designated to be 'grammatical' and other regularities deemed to be 'rhetorical', 'formulaic', etc.

2) Because grammar is always emergent but never present, it could be said that it never exists as such, but is always coming into being. There is, in other words, no 'grammar' but only 'grammaticization'— movements toward structure which are often characterizable in typical ways. It goes without saying that many phenomena which we would agree to call grammatical are relatively stable and uniform. That is not in dispute. The point again is that any decision to limit the domain of grammar to just those phenomena which are relatively fixed and stable seems arbitrary.

(3) The major descriptive project of Emergent Grammar is to identify recurrent strategies for building discourses — strategies which have intra-linguistic or inter-linguistic generality (or both) and which move toward grammaticalization along parallel lines.

In studying discourse with a view to describing emergent regularities, it is therefore most useful to begin by establishing frequently occurring, relatively stable clause types. A useful concept here is that of the 'figure', suggested by Pete Becker. A figure is a phrase or clause which is highly standardized in its format and which permits substitution in a few restricted places. It has a rudimentary internal structure, but it is much closer to a formula than to freely generated 'sentences'. To the extent that discourse is not prefabricated, it consists for the most part of assemblages of a small number of such figures. Knud Lambrecht's notion of a 'Preferred Clause Unit' seems to be quite similar, only Becker's concept of a 'figure' permits a number of such types of clause unit to be reckoned with. Consider the following examples from Old English (Plummer 1899):

1.     ond *ða* geascode he *ð*one cyning 'And then he found the king'  
       lytle werode 'with a small band of men'  
       on wifcuð*ð*e 'a-wenching'  
       on Merantune, 'in Merton'  
       ond hine *ð*ær berad, 'and caught up with him there'  
       ond *ð*one bur utan be eode 'and surrounded the hut outside'  
       *ær* hine *ða* men onfunden 'before the men were aware of him'  
       *ð*e mid *ð*am kyninge wærun; 'who were with the king'
  
- 2 *ða* ridon hie *ð*ider, 'Then they rode up'  
       ond his aldormon Osric, 'and his alderman Osric'  
       ond Wifer*ð* his *ð*egn, 'and his thane Wiferth'  
       ond *ða* men 'and the men'  
       *ð*e he be *æ*ftan him læfde *ær*, 'which he left behind him earlier'  
       ond *ð*one *æ*ðeling on *ð*ære byrig metton, 'and met the prince in  
the villa'  
       *ð*ær se cyning ofslægen læg... "where the king lay slain"

'Then his alderman Osric, his thane Wiferth, and the men he had left back earlier rode up, and found the prince in the compound where the king lay slain...' (755 AD)

Here, a handy way of building up a discourse, such as a narrative, is to construct it by means of a verb-initial clause, usually preceded by a temporal adverb such as *ðā* 'then'; this clause typically elaborates a setting for an action, and may contain a number of lexical nouns introducing circumstances and participants:

ond *ða* geascode he *ðone* cyning 'And then he found the king  
lytle werode 'with a small band of men  
on wifcuððe 'a-wenching'  
on Merantune, 'in Merton'

*ða* ridon hie *ðider*, 'Then they rode up'  
ond his aldormon Osric, 'and his alderman Osric'  
ond Wiferð his *ðegn*, 'and his thane Wiferth'  
ond *ða* men 'and the men'

It is followed by a succession of verb-final clauses, in which lexical NP's are minimally represented. These verb final clauses are built up of a particle such as *ond* 'and', one or more initial pronouns unrestricted as to case, perhaps a lexical noun or adverb, and then the verb:

ond hine *ðær* berad, 'and caught up with him there'  
ond *ðone* bur utan be eode 'and surrounded the hut outside'  
*ær* hine *ða* men onfunden 'before the men were aware of him'  
*ðe* mid *ðam* kyninge wærun; 'who were with the king'

It is not a question of an invariant hyperform from which different clauses are derived by processes of deletion and movement. Instead it seems that constructions spread outwards from a small nucleus and in turn form new nuclei (something like the metastasis of malignant cells, to co-opt a metaphor of Bolinger's), and the resultant array of clauses are in 'family resemblance' relationships to one another. Among the features of the figure is a tendency to avoid phrasing which would lead to multiple lexical noun participants in the same figure. So in (2), the multiple agents 'they', 'his alderman Osric', 'his thane Wiferth' and 'the men' are distributed over several phrases, and only one of them (hie 'they') is retained in the figure. Lambrecht's work on Spoken French (e.g., Lambrecht, 1987) shows how a large part of its grammar is invested in strategies for preserving the external form of what he calls the Preferred Clause Unit.

### 3 The Malay Ergative: An Emergent Construction.

I will conclude by considering some consequences of emergent grammar for morphology. A major postulate, or working hypothesis, of Emergent Grammar is that the more useful a construction is, the more it will tend to become structuralized, in the sense of achieving

cross-textual consistency, and serving as a basis for variation and extension. An elementary example of this is 'Watkins' Law' (Watkins 1962: 93-96; Collinge 1985: 239-240). Calvert Watkins has noted that the third person singular of a paradigm forms the basis for new paradigms. The particular interest of Watkins' observation is his point that there are asymmetries among the persons, which in fact play quite different roles in discourse, having eventual consequences for the development of paradigms.

In written Malay texts (those used in this study were Abdullah 1932, Abdullah 1928, both written in the 1840's), a highly frequent and favored clause type consists of a transitive verb with an enclitic ergative pronoun, followed by a simple lexical patient. This clause type is found in numerous types of context; the following exemplify narrative (3), and procedural (4) discourse; the verb + clitic complex is italicized:

- (3) Hata maka *di-panggil-nya* aku masok ka dalam bilek 'After that he summoned me into the room'  
 tempat ia menulis, 'where he wrote'  
 maka *di-tulis-nya* sa keping surat; 'and he wrote a letter;'  
 sa telah sudah, 'when he had finished'  
 maka *di-buka-nya* peti-nya, 'he opened his sea-chest'  
*di-ambil-nya* tiga puluh ringgit, 'took out thirty dollars'  
*di-unjokkan-nya* surat serta wang itu, 'handed [me] the letter and the money'
- (4) Maka erti salang itu, 'And [execution by] salang means'  
*di-ikat-nya* kaki tangan orang itu, 'they bind the man hand and foot'  
 lalu *di-dudokkan-nya* di-haluan perahu, 'and put him in the bow of a boat'  
*di-kayohkan-nya* kapada sa buah anak sungai. 'and row him to one of the backwaters of the river'

The characteristic particles *lalu* and *maka* are closely comparable to the *ond* of Old English; the verb has a prefix *di-*, which serves as an agreement prefix with third person agents, and usually an enclitic *-nya* meaning 'he, they (ergative)', giving a transitive clause beginning with *di-V-nya* 'he, they V'ed (it)'. Discourses may now be constructed by stringing together these transitive clauses, together with a few other quite easily characterizable types, substituting new nouns and verbs as needed, but generally keeping the basic shape of the figure intact.

The argument structure of these figures is along the lines of Du Bois' Preferred Argument Structure (Du Bois 1986). Agents are generally continuous as topics, and are either zeroed or represented simply by the enclitic *nya*. Lexical nouns are for the most part non-agents, such as patients, indirect objects, and obliques of various kinds. If there is a lexical agent, this has a preposition *oleh*, 'by'. But it will be noticed that lexical agents are relatively few and far between; some examples of them are:

(5) maka anak-nya perempuan itu pun hendak menangkap ikan itu,  
'and his daughter tried to pick up the fish'  
sa-telah di-tangkap-nya dari ekor-nya, 'when she took hold of it by its tail'  
maka di-kebaskan oleh [ERG.] ikan itu tangan-nya, "the fish jolted her  
hand'

(6) maka anjing itu hendak pergi menchari ayer di-sungai itu, 'and  
the dog went down to the river to find water,'  
maka tiba-tiba di-sembar oleh [ERG] buaya 'and was suddenly snapped  
up by a crocodile.'

(7) Maka oleh [ERG] Tuan Farquhar 'And Mr. Farquhar'  
di-suroh-nya ambil bangkai buaya itu, 'had them get the crocodile's  
body,'  
di-gantong-nya di-pohon jawi-jawi 'and he hung it from a fig-tree'

In example (5), there are two lexical nouns, an agent *oleh ikan itu* 'by the fish', and an absolutive, *tangan nya* 'her hand', and the verb is *di-kebaskan* 'shock, jolt [her]'. In example (6), the agentive phrase is *oleh buaya* 'by a crocodile', and the patient is zero, being continued from the previous clause; the verb is *di-sembar* 'snap [it] up'. Transitive agents which are lexical nouns, such as *buaya* 'crocodile' in (6) above, take the preposition *oleh* provided they are specific participants in the discourse. 'Specific' usually means definite in the sense of having been referred to previously in the discourse; but the noun may be new, as here, and its individuation then depends on subsequent mentions in the discourse. Lexical agents which are neither old nor subsequently mentioned — i.e., which do not qualify as specific participants in the discourse — do not take *oleh*, as in:

(8) ada yang di-makan harimau 'some were eaten by tigers'

where *harimau* 'tiger' is the lexical agent of *di-makan* 'eat (transitive)'. These 'indefinite, non-specific' nouns, then, behave like agentive pronouns in lacking the preposition and being placed immediately adjacent to the verb stem.

Now the presence of an ergative preposition before a lexical agent is explained by the Preferred Argument Structure. Lexical agents, being highly marked in discourse terms, must receive a special indicator, in this case the agentive preposition *oleh*. But indefinite-nonspecific lexical agents like *harimau* 'tiger' in (8) should not be exempt from case marking. On the contrary, they are if anything even more highly marked as agents than definite specific lexical agents. I return to this point in (iii) below.

Although most transitive agents are placed immediately after the verb in the preferred clause unit, the lexical transitive agent may also appear outside the clause, in very much the same way that we saw in Old English that 'extra' NP's are placed outside the nuclear clause. Thus in the next example the ergative agent is placed outside the clause in front of the verb:

(9) Maka oleh [ERG] Tuan Farquhar 'And Mr. Farquhar'

di-suroh-nya ambil bangkai buaya itu, 'had them get the crocodile's  
body,'  
di-gantong-nya di-pohon jawi-jawi 'and he hung it from a fig tree'

There are numerous examples of this 'extraposition' of the ergative; it is especially found when a single lexical agent is shared by several subsequent clauses, as here. The ergative phrase then has a domain which extends over a number of clauses, and in fact has features of an independent clause in its own right.

Now I want to suggest that this is exactly what is happening— that the prepositional (lexical) ergative is emerging out of a 'serial verb' construction which sometimes re-appears in its original clausal form, in much the same way that the English indefinite article sometimes appears in contexts where its earlier specific sense is reflected. My reasons for saying this are the following:

(i) That the 'preposition' *oleh* is verbal in origin is indisputable. Compound forms of the verb still exist: *beroleh* 'to obtain', *oleh-oleh* 'something brought back as a gift', and the modal *boleh* 'be able, be allowed to'. They suggest a meaning like 'acquire, achieve, manage, accomplish' which seems well within the typology of grammaticization of ergative prepositions out of verbs.

(ii) The possible independence of the agentive clause from the action clause is seen nicely in the next example (10), in which the verb is in the *meng-* prefixed form rather than the *di-* form ('passive') otherwise invariably found with the ergative:

(10) Maka oleh Grandpre memberikan-lah surat itu ka tangan Enche Ha  
and [name] meng:give PCLE letter the to hand Mr. [name]  
'And Grandpre handed the letter over to Mr. Ha.'

The sense that the argument 'Grandpre' is shared between two clauses in a serial fashion is striking.

(iii) As previously noted, the agent of the ergative with *oleh* is always definite or specific. In other words, the lexical agent with *oleh* retains characteristics of a topic/agent, and no doubt reflects an original definiteness constraint on agent/topics. By contrast, lexical agents which are non-specific were never appropriate topics of *oleh*, while non-lexical (i.e., pronominal) agents were always cliticized to the main verb.

(iv) When the ergative agent is separated from the action clause, the action clause usually also has the clitic agentive *-nya*, that is, the agent is referred to twice, as in example (9). This is exactly the same as when an agent is introduced in a previous separate clause and referred to again.

This is of course an example of grammaticization of the classical kind which has often been noticed in the literature. What I have wanted to stress here is the need to understand not only the formal process but the way in which that formal process emerges from a discourse context, in other words, is anchored in particular, concrete utterances. It is this 'prior textuality' of the construction which explains why it has retained properties of a separate, external clause. What we see emerging, then, is a new strategy for permitting a

lexical agent to be incorporated into a nuclear clause under certain contextual conditions, presumably involving differences of topic continuity.

#### 4. Conclusion

I conclude this paper with some syllogisms, extrapolated from the first couple of pages of Radford's textbook on transformational syntax. References, with emphasis as in the original, are to Radford 1981:

“What is a grammar of a language? Chomsky gives an essentially mentalist answer to this question: for him a grammar is a model (=systematic description) of those linguistic abilities of the native speaker of a language which enable him to speak and understand this language fluently. ... Thus a grammar of a language is a model of the linguistic competence of the fluent native speaker of the language.” (p. 2)

“...in the case of a sentence such as:

(1) He thinks that John is wrong

it is the native speaker's grammatical competence (his knowledge of the grammar of his language) which tells him that *he* cannot be interpreted as referring to the same person as *John* in (1).” (p. 3)

#### *Some Syllogisms*

[1A] A grammar of a language is a model of the linguistic competence of the fluent native speaker of the language.

[1B] A model is a systematic description.

THEREFORE:

[1C] A grammar of a language is a systematic description of the linguistic competence of the fluent native speaker of the language.

[2A] Grammatical competence is the native speaker's knowledge of the grammar of his language.

[2B] (= [1C]) A grammar of a language is a systematic description of the linguistic competence of the fluent native speaker of the language.

THEREFORE:

[2C] Grammatical competence is the native speaker's knowledge of a systematic description of the linguistic competence of the fluent native speaker of the language.

[3A] “...in the case of a sentence such as:

(1) He thinks that John is wrong

it is the native speaker's grammatical competence (his knowledge of the grammar of his language) which tells him that *he* cannot be interpreted as referring to the same person as *John* in (1)." (p. 3)

[3B] = [2C] Grammatical competence is the native speaker's knowledge of a systematic description of the linguistic competence of the fluent native speaker of the language.

THEREFORE:

[3C] ... in the case of a sentence such as:

(1) He thinks that John is wrong

it is the native speaker's knowledge of a systematic description of the linguistic competence of a fluent native speaker of the language which tells him that *he* cannot be interpreted as referring to the same person as *John* in (1).

It will be seen that 'grammar' begins life on page 2 in its theoretically correct style, as a 'model' of the native speaker's 'linguistic competence'. But notice that by page 3, 'grammar' is suddenly no longer a linguist's construct, a formal characterization of the abilities presumed to underlie the speaker's behavior, but the knowledge itself. It has gone from a linguist's theory to something the speaker possesses. One would not blame Radford, were it not that formal grammarians are quick to castigate discourse linguists for alleged 'confusion' over the notion of 'grammar', and often accuse them of not understanding this supposedly elementary concept.

There is no question that 'grammar' is an infuriatingly elusive notion, and that it is very easy to have a clear idea about what 'grammar' is in the sense of being able to give an abstract definition of it, but quite another to apply that definition consistently in practice. This asymmetry suggests that the notion of grammar is intrinsically unstable and indeterminate, relative to the observer, to those involved in the speech situation, and to the particular set of phenomena being focused upon. It suggests also that we need to question the supposition of a mentally represented set of rules, and to set aside as well the idea in Fromkin's statement which I quoted earlier, that speakers possess an abstract linguistic system ready and waiting to be drawn upon — 'accessed'! — in case they should ever need to speak.

## References

- Abdullah ibn Abdul Qadir. 1932. *Hikayat Abdullah*. Singapore: Malaya Publishing House. *Malay Literature Series 4*.
- Abdullah ibn Abdul Qadir. 1928. *Kesah Pelayaran Abdullah*. Singapore: Malaya Publishing House. *Malay Literature Series 2*.
- Bakhtin, M. M. 1981. "Discourse in the novel." In *The Dialogic Imagination: Four Essays*, edited by Michael Holquist, translated by Carl Emerson and Michael Holquist. Austin: University of Texas Press.



- Becker, Alton L. 1979. 'The figure a sentence makes: An interpretation of a Classical Malay sentence.' In Talmy Givon, ed., *Discourse and Syntax*, 243-260. Academic Press. (*Syntax and Semantics* 12)
- Bolinger, Dwight. 1976. 'Meaning and memory'. *Forum Linguisticum* 1,1: 1-14.
- Bosworth, Joseph, and T. Northgate Toller. 1898. *An Anglo-Saxon Dictionary*. Oxford: Oxford University Press.
- Clifford, James. 1986. 'Introduction: Partial Truths.' James Clifford and George Marcus, eds., *Writing Culture. The Poetics and Politics of Ethnography*, 1-26. Berkeley: University of California Press.
- Collinge, N. E. 1985. *The Laws of Indo-European*. Amsterdam: John Benjamins BV. (Current Issues in Linguistic Theory, vol. 35)
- Culler, Jonathan. 1982. *On Deconstruction: Theory and Criticism after Structuralism*. Ithaca: Cornell University Press.
- Du Bois, John. 1986. 'Competing motivations.' In John Haiman, ed., *Iconicity in Language*. Amsterdam: John Benjamins BV. (Typological Studies in Language)
- Fromkin, Victoria. 1985. 'Linguistics', in: 'Major Trends in Research: 22 Leading Scholars Report on Their Fields.' *The Chronicle of Higher Education* Sept. 4, 1985, pp. 12-13.
- Hopper, Paul J., and Janice Martin. 1987. 'Structuralism and Diachrony: the development of the indefinite article in English.' in Ramat, Anna Giacalone, Onofrio Carruba, and Guiliano Bernini, eds. *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: Benjamins.
- Lambrecht, Knud. 1984. 'Formulaicity, frame semantics, and pragmatics in German binomial expressions.' *Language* 60,4: 753-796.
- Lambrecht, Knud. 1987. 'On the status of SVO sentences in French discourse.' In R. Tomlin, ed., *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins BV.
- Morgan, Jerry. 1981. 'Some observations on discourse and sentence grammar.' *Studies in the Linguistic Sciences* 11: 137-144.
- Newmeyer, Frederick. 1983. *Grammatical theory: its limits and possibilities*. Chicago: University of Chicago Press.
- Plummer, Charles, ed. 1892. *Two of the Saxon Chronicles, Parallel. ... Edited on the basis of an edition by John Earle*. Oxford: Clarendon Press.
- Radford, Andrew. 1981. *Transformational Syntax: A Student's Guide to Chomsky's Extended Standard Theory*. Cambridge University Press. (Cambridge Textbooks in Linguistics)
- Sadock, Jerry. 1984. 'Whither radical pragmatics?' In D. Tannen, ed., Georgetown University Round Table Conference on Linguistics, 139-149.
- Smith, Barbara Harnstein. 1978. *On the Margins of Discourse*. University of Chicago Press.
- Staten, Henry. 1984. *Wittgenstein and Derrida*. University of Nebraska Press.
- Watkins, Calvert. 1962. *Indo-European Origins of the Celtic Verb*. Dublin: Institute of Advanced Studies.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Trans. G. E. Anscombe. 3rd edition. New York: Macmillan.

## Footnotes

<sup>1</sup>I would like to thank Sandy Thompson for written comments on the conference version of this paper. It goes without saying that numerous others have directly or indirectly influenced the paper, but its general debt to Sandy and to Pete Becker, Ranjit Chatterjee, Jack Du Bois, Barbara Fox, Talmy Givon, Knud Lambrecht, Larry Roberts, Deborah Tannen, and especially Dwight Dolinger, is surely both pervasive and obvious; none of them should be blamed for any errors or excesses in the use I have made of their ideas, and I apologise to them if, because of space restrictions, I have not always cited their published work when I might have.

<sup>2</sup>This was pointed out to me independently by Catherine Lutz and Deborah Tannen.

<sup>3</sup>Again, Derrida's proposal for a typology of 'grafts' seems closely relevant here.