# Competing motivations and emergence: explaining implicational hierarchies

Simon Kirby
simon@ling.ed.ac.uk
*To appear in Language Typology – 1997*

**Abstract**

It is the basic tenet of the functional approach to typology that at least some linguistic universals may be explained by appealing to features of language use. But the mechanics of the mapping between function and distribution are seldom made explicit. In this paper, a theory of linguistic adaptation is set out in which universals are treated as *phenomena of the third kind* in Keller's (1994) terms, i.e. objects which are the result of human actions but not the goal of their intentions. In this view universals emerge at a global level from the interactions of individuals whose actions locally can be described in terms of functional motivations. Computational models of this process (an *invisible hand* process) can be set up to see what universals emerge with different functional pressures in place.

Using this model, Keenan and Comrie's accessibility hierarchy is shown to emerge when competing functional pressures relating to structural complexity and morphological complexity are taken into account. This is contrasted with claims that structural complexity asymmetries *alone* can explain implicational universals. The value of this approach is further demonstrated by testing the predicted skewing in the distribution of 'case-coding' strategies down the hierarchy.

## 1 Introduction

A central goal of theoretical linguistics is the explanation of the linguistic universals revealed by typological research. The types of explanations sought fall into a range of categories (see, e.g. Hawkins 1988; Hurford 1990; Hawkins 1992) and it is unclear whether there are any underlying principles in common with the various approaches. This paper is part of an attempt to provide a theoretical framework for various kinds of functional-typological explanations that is also broadly compatible with the nativist approach to language acquisition (Kirby 1994b; Kirby 1996).

The approach relies crucially on the idea of selective adaptation:

> "...the sorts of explanations made by typologists are essentially adaptive ones: language structures are the way they are because of their adaptation to the function(s) of language ... In this respect linguistics also parallels biology.
>
> However, the philosophical analogy between linguistic functional explanations and biological adaptation is not always fully worked out in linguistics." (Croft 1993)

The framework put forward in this paper is an explicit description of the adaptive process. In other words it is an attempt to solve the problem stated below (also discussed in other forms by Hall 1988 and Bybee 1988):

(1) **The problem of linkage:** How does a universal property of language use give rise to a restriction on the distribution of occurring languages in the space of possible languages?

For the purposes of this paper, the linkage in question is that between properties of innate language processing mechanisms and *hierarchical universals*. These are chains of implicational universals of the form $(P_1 \rightarrow P_2)\&(P_2 \rightarrow P_3)\&\ldots\&(P_{n-1} \rightarrow Pn)$ commonly written $P_n > P_{n-1} > \ldots > P_3 > P_2 > P_1$, where $P_i$ is associated with some binary property whose value can be determined for all languages. Of course, these hierarchies predict that only the following language types will occur (where + indicates that the language has the property $P_i$, and – indicates that the language has the converse property $P_i'$):

(2)

| $P_n$ | $P_{n-1}$ | $\ldots$ | $P_3$ | $P_2$ | $P_1$ |
|---|---|---|---|---|---|
| − | − | ... | − | − | − |
| + | − | ... | − | − | − |
| + | + | ... | − | − | − |
| ⋮ | | | | | |
| + | + | ... | + | − | − |
| + | + | ... | + | + | − |
| + | + | ... | + | + | + |

The particular universal that this paper examines was reported by Keenan & Comrie (1977). They show that the accessibility of noun phrases to relativisation depends on the grammatical function of the gap or resumptive pronoun within the relative clause:

(3) Subject>Direct Object>Indirect Object>Oblique>Genitive>Object of Comparison

Hawkins' (1994) explanation of this universal relies on a claim that the ease of parsing of these relative clause constructions decreases down the hierarchy in (3), and that this leads to the implicational constraints on cross-linguistic distribution. The intuition is that languages somehow select a point on a hierarchy of processing complexity below which relative clauses will be grammatical, and above which they will be ungrammatical. Indeed, he uses a similar argument to explain a range of hierarchies:

> "Implicational hierarchies of center embeddings ..., of relative clause formations within Relativisation Domains, and of WH-extractions within Movement Domains are all syntactic complexity rankings explainable in terms of relative degrees of processing and efficiency.... These hierarchies define the sequence in which grammatical variants are selected within each grammatical domain, and the claim is being made that this sequence involves increasing complexity, and that the cut-off points represent a conventionalized response by speakers of each language not to tolerate processing difficulty or inefficiency below that point."(Hawkins 1994:435)

Notice, that, if the complexity referred to is related to parsing, then speakers must be responding in some way to the needs of hearers in their tolerance of the relevant structures. I will refer to this view as the *assumption of speaker altruism*.

Although Hawkins' description of structural complexity captures an important generalisation about the pressures constructions pose for processing, I will show that this is not quite enough to explain the origin of the cross-linguistic hierarchy. Within the model that is outlined here, any explanation that relies solely on a gradient measure of complexity to explain an implicational universal will fail.

The main claims that this paper makes are:

- Explanation of cross-linguistic hierarchies must appeal to *competing motivations*.
- Appealing to functional pressures in conflict does not lead to vacuous explanations.
- Global structures (such as hierarchy) emerge from local interactions.
- The assumption of speaker altruism is not required.
- The mechanism of explanation should be made explicit within a mathematical or computational model.

## 2 Structural complexity

The hierarchy in (3) is examined in some length by Keenan & Comrie (1977); Comrie & Keenan (1979); Keenan & Comrie (1979); Maxwell (1979) with respect to relative clauses (RCs). Although the hierarchy is assumed to have wider application (e.g. causatives), the explanation discussed only deals with RCs. The relation of the hierarchy to RC phenomena is given by a set of definitions and constraints:

(4) **Subject relative universal:** "All languages can relativise subjects." (Comrie & Keenan 1979:652) (A strategy that can relativise subjects is a *primary strategy*.)

(5) **Accessibility hierarchy constraints:**

a. "If a language can relativise any position on the AH with a primary strategy, then it can relativise all higher positions with that strategy.

b. For each position on the AH, there are possible languages which can relativise that position with a primary strategy, but cannot relativise any lower position with that strategy" (Comrie & Keenan 1979:653)

Keenan & Hawkins (1987) report results from a psycholinguistic experiment testing native English speakers' 'mastery' of relative clauses down the AH. The experiments were designed to test repetition of RCs that occurred modifying subjects in the matrix clause, so no conclusions can be drawn about: a) other languages, b) RCs modifying matrix objects etc., or c) whether the AH affects production, or perception, or both. These points aside, the mastery of RCs clearly declined down the AH. As Keenan and Hawkins point out, this processing difficulty might explain the AH. Other experiments have been carried out that have tested the relative processing difficulty of RCs on the first two positions on the hierarchy (subject and direct object). MacWhinney & Pleh (1988) review a number of studies in comprehension in English children that are consistent with the view that subject relatives are easier to parse than object relatives (though see Kirby 1996 for further discussion). Furthermore, their own study of Hungarian reveals a similar pattern.

Hawkins (1994) claims that these results are independently derivable from a theory of structural complexity that can in turn be used to explain a range of universals, including the AH. His theory defines a measure of tree-complexity associated with a particular node in a constituent that is relative to *a particular psycholinguistic operation*. The complexity of processing a relative clause is proportional to the size of a portion of the tree that is involved in co-indexing the trace, or pronoun, in the clause with its head noun. Hawkins' definitions (pp. 28–31) cash out as follows:

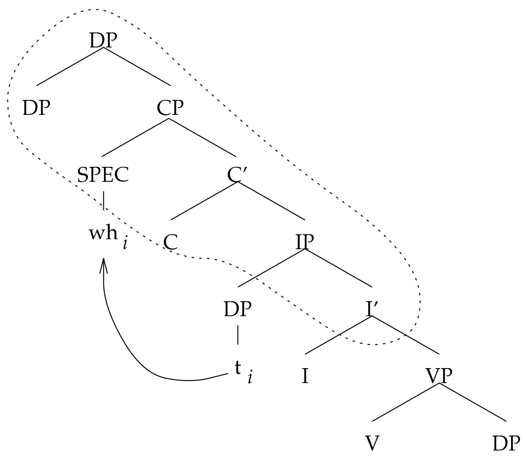
(6) **Structural complexity of relative clause:** The structural complexity of a relative clause is calculated by counting the following nodes:

- all nodes dominating the trace, or pronoun, within the NP dominating the RC (including the NP itself),
- the sisters of the trace or pronoun,[1]
- the sisters of the nodes dominating the trace, or pronoun, within the NP dominating the RC.

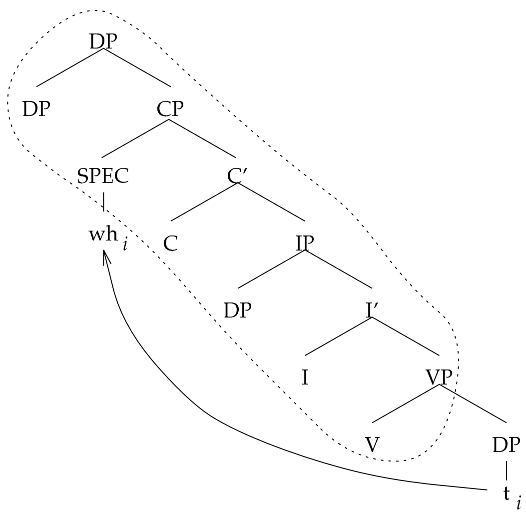
The intuition captured by this definition is that relating the head noun with a trace (or pronoun) becomes more complex the more the trace (or pronoun) is embedded within the subordinate clause.

Hawkins demonstrates this metric using tree structures that rely on traditional notions of constituency, but the complexity rankings seem to remain the same if they are calculated using other syntactic analyses. Consider the structures below, which are standard treatments of relative clauses within the Principles and Parameters tradition:

(7)




(8)



The first tree (7) is a structure where the subject DP in [Spec,IP] has moved to [Spec,CP]. This then, is the structure of a subject relative. The nodes that are involved in the calculation of complexity as defined by (6) are circled. The second tree (8) is the equivalent tree for an object relative – in this case it is obvious that the RC-complexity is higher. Similar arguments can be made for the relative ranking of other positions on the AH (Hawkins 1994:39–41).

This account is successful inasmuch as it predicts the relative ranking of relative clauses in a hierarchy of parsing complexity, and uses concepts – such as structural domains – that can be generalised to other domains (e.g. word order and extraction). However, the theory as it stands does not answer the problem of linkage stated in (1); exactly how do the structural complexity facts end up being expressed cross-linguistically? The next section attempts to answer this question in such a way as to provide a general solution to the problem of linkage.

## 3 Selection and emergence

Keller (1994) puts forward an *invisible hand* account of language change. In this theory, language changes are viewed as *phenomena of the third kind*. Essentially, Keller gives us a typology of phenomena, dividing explananda into natural phenomena and results of human action, and further dividing the latter into artifacts and phenomena of the third kind. These phenomena can be characterised as those "things which are the result of human actions but not the goal of their intentions" (Keller 1994:56). The process that gives rise to these phenomena is termed the invisible hand process.

### 3.1 Universals are phenomena of the third kind

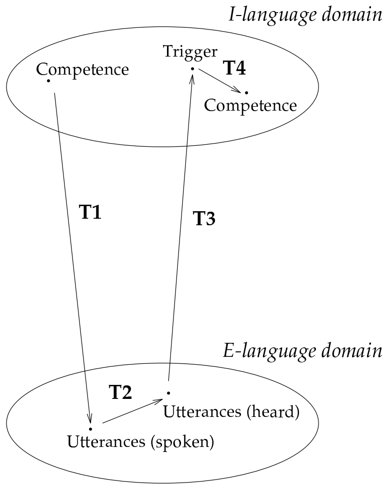
Keller discusses individual language changes as instances of objects of this kind. Keller points out that the explanation for the phenomenon of language change must refer to the actions of users of the language, and yet particular changes cannot be said to have been their *goal*. The *invisible hand process* describes the non-local consequences of users behaviour (see Keller 1994, 93–95 for an example of this type of explanation). Particular changes in this view are *emergent properties* of the interaction of users during those changes.

Universals are similarly non-intentional results of human action. In other words, the local, individual actions of many speakers, hearers and acquirers of language across time and space conspire to produce non-local, universal patterns of variation. A description of the invisible hand process in this case is a theory of the propagation of variation through individuals. Indeed, the same mechanisms that explain individual language changes can be called upon to explain universals. A particular universal such as the AH can be thought of as a *higher order emergent property*.

This brief discussion points to some desirable features we might look for in an explanation for universals. In particular, we should hope to only make reference to the actions of individuals at individual points in time. Furthermore, our model of the individual must describe precisely the relationship between these actions and the ecological conditions in which the individual is situated.

### 3.2 Linguistic selection

(9)




Kirby (1994a); Kirby (1996) argues that for an explanation of the desired type to work, the influence of processing on language competence must be seen as a selective influence. More properly, functional pressures must influence the selection of linguistic variants that are *competing* in some way, and this selection must occur at some point in the cycle of language acquisition and use. There is some transformation that maps the competence of a speaker at time $t_1$ to the competence of a speaker in the same speech community at some later time $t_2$. Functional selection influences this transformation in a predictable, though statistical, manner.

Viewing linguistic evolution in terms of laws of transformation closely parallels biological thinking. So closely, in fact, that we can usefully borrow a map of transformations from Lewontin (1974) (cited in Sober 1984), replacing *genotypes* with *I-language* and *phenotypes* with *E-language* as in diagram (9). The transformation from competence to competence involves objects in two very different domains. *I-language* (or internalised-language) consists of objects in individual speakers' brains. These objects, the domain in which they exist and the transformation **T4** (acquisition), are what Chomsky (1986) argues are the proper target of study in linguistics.

On the other hand we have *E-language* (or externalised-language) which consists of *utterances* in some broad sense. These objects are more ephemeral, and are typically viewed as epiphenomena in the Chomskyan program. The transformation **T2** involves features of the world at particular points of time, for example, the level of noise, the availability of hearers, and so on.

Finally, we have the transformations **T1** and **T3** which map objects in one domain to those in the other. The former is mediated by speakers (production), and the latter by hearers (parsing). Both these transformations and those that map between objects within domains, are not well understood by linguistic theory, but it is generally assumed that some innate (and therefore universally shared) neurological mechanisms play a role. In particular the focus of the explanation in this paper is on the role of complexity of processing in influencing the transformations **T1** and **T3**.

## 4 The computational model

A simulation embodying the model discussed above has been built in order to test what universals may emerge given some assumptions about complexity. The simulation approach is useful since it forces one to make explicit tacit assumptions about the solution to the problem of linkage (1). As will be seen, this process of making explicit itself is important since it shows that Hawkins' complexity metric described in section 2 above does *not* on its own give rise to the expected result.

The simulation discussed in this paper examines only the first two positions on the accessibility hierarchy – subject and direct object. Discussion in later sections shows how the results are easily extended to the rest of the hierarchy, and provide an explanation for the subject relative universal (4) that we will ignore for the moment. The relevant components of the simulation are:

1. A 'space' of language users arranged as a 40x40 square. Each language user is in one of four states $SO, S'O, SO', S'O'$. $Sx$ means that the user has an I-language that allows subject relatives, $xO$ means that the user has direct object relatives. The variants with the prime mean that the user only has access to a non-relativised equivalent. So, for example, a speaker of English would be represented with the type $SO$, whereas a speaker of Maori (which cannot relativise on direct objects) would be represented with the type $SO'$.

2. Each user has four 'neighbours' which are directly above, below, to the left, or to the right of it. (The space is toroidal, so that a user on the left edge has a neighbour on the far right, and a user on the top has a neighbour on the bottom.)

3. Each generation, all the language users produce $n$ utterances in line with their I-language states randomly. (This is a trivial implementation of transformation **T1** that maps I-languages states directly onto utterances – later we shall see that this step needs to be modified.)

4. Each generation, a new space is created for the next generation of users. These users acquire their I-language states taking five of the last generation's users' utterances as input. The five users involved are: the one in the same place as the new user, and the four neighbours of that user.

   The acquisition of the new state must reflect the distribution of variants that the acquirer hears, but it must also take into account the complexity of the variants. This is done by counting the numbers of the variants in the input, and then forming a probability of acquisition for each variant by weighting the numbers towards those that are easier to parse. In other words:

   - Pool all the utterances together (transformation **T2**) and, for each variant $v$, acquire that variant with a (normalised) probability that reflects the number of that variant in the input $n_v$ and the relative structural complexity of the variant $c_v$ (which is set up to vary between 0 and 1). This implements transformation **T3** by probabilistically filtering out the utterances that are harder to parse, and trivially implementing **T4** by mapping utterances onto I-language states. The probability of acquiring $v$ will be:

$$p(v) = \frac{n_v(1-c_v)}{n_v(1-c_v) + n_{v'}(1-c_{v'})}$$

5. The old space is discarded, and the new space becomes the new generation.

The first feature of the simulation results, which is largely independent of the initial conditions, is that large groups of similar individuals – language 'communities' – quickly form. This is a similar result to one of Jules Levin (reported in Keller 1994, 100). Levin's simulation is similar to this one in many respects, but it does not model the influence of the transformations **T1** and **T3**, the transformations that map between I-language and E-language. In other words, it assumes that the language that an individual will acquire is simply the one that most of that individuals neighbours has.[2] Keller (1994:99) calls this 'Humboldt's Maxim':

> "Talk in a way in which you believe the other would talk if he or she would talk in your place. My thesis is that this maxim – a slightly modified version of Humboldt's own formulation of it – produces homogeneity if the starting point is heterogeneous and stasis if the starting point is homogeneous."

Indeed, this is what happens with Levin's simulation. Starting with a random patterning of two types, the simulation finally settles down with the types clustering together in large groups.

What of the final state of the simulation in this paper? The expected result, if Hawkins' theory can be simply 'plugged-in' to an invisible hand explanation, is that the $SO$,$SO'$ and $S'O'$ types should survive, but that $S'O$ type should not. This is the prediction of the first part of the AH: Direct Object→Subject. (Recall that the explanation of the separate subject relative universal (4) is to be dealt with later.) However, as long as the structural complexity of relative clauses is higher than their non-relative variants ($c_S > 0.5$ and $c_O > 0.5$) the end result is just languages of type $S'O'$. A typical result is shown in figure 1. The different language types are represented by differing shades of grey. Notice that from an initial random condition the type corresponding to $S'O'$ soon predominates. The relative ranking of the complexity of subject relatives and object relatives makes no difference to this result. If, on the other hand, the complexity of relative clauses is assumed to be *lower* than the non-relative variants ($c_S < 0.5$ and $c_O < 0.5$) – an odd assumption given Keenan & Hawkins (1987) results, in any case – then the final state is languages of type $SO$ only, again the relative ranking of $S$ and $O$ complexity is irrelevant.

## 5 Competing motivations

The solution to this problem involves a 'competing motivations' explanation (e.g. Hall 1992; DuBois 1987; Givón 1979). These are explanations that rely on functional pressures in *conflict*. Newmeyer (1994) examines several different types of these explanations and argues that *some* attempts by functionalists to build these sorts of motivations directly into their theories of synchronic grammatical phenomena render both their descriptions and their explanations inadequate. These criticisms will not apply to the approach taken in this paper since the functional pressures in question are not assumed to be encoded in grammars. Instead, the I-language domain is taken to be autonomous from the environment; however, as the model

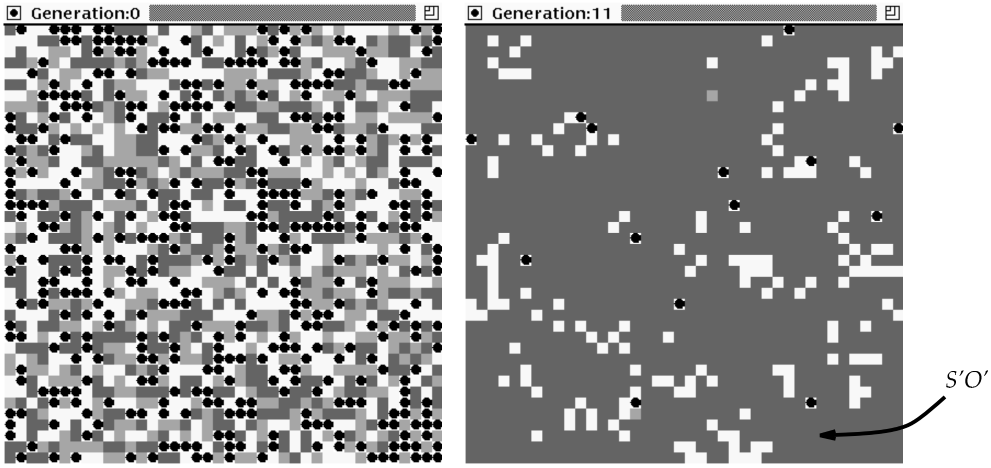



Figure 1: An example run of the simulation showing predominance of $S'O'$.

described in the previous section makes clear, this does not preclude the possibility that functional pressures can influence the possible states that a grammar can take.

The influence on parsing of structural complexity is *one* functional pressure that affects relative clauses. Because it affects parsing, it is part of what I will call *p-complexity*. The details of a full definition of p-complexity will involve many different aspects, but the influence of it within the selection model is simple:

(10) **p-complexity:** In comprehension, the selection of competing variants (i.e. variant forms that are synonymous, or functionally undifferentiated) will depend on their relative parsing complexity. So, the more difficult some variant is to parse, the more likely it will fail to be included in the set of trigger experiences of the child.

Some of the other factors that influence p-complexity are, for example, redundancy of information (see section 8), and configurational markedness (see Kirby 1994a). For the moment it suffices to see p-complexity as identical to Hawkins' metric given in (6).

Another type of complexity that will influence the selection model is morphological or *m-complexity*:

(11) **m-complexity:** In production, the selection of variants will depend on their relative morphological complexity. So, given two competing ways in which to produce some message, the speaker will be more likely to produce the one that is less morphologically complex.

Traditional structural markedness, where a marked form has more morphemes (e.g. Croft 1990, 73), is clearly related to m-complexity. However, precisely how this affects production is not clear: is the relevant measure the number of morphemes, or the number of morphs? Do all morphemes carry equal m-complexity, or are morphemes that are involved in agreement ($\phi$-features) more complex to produce than others (such as definiteness markers)? We shall return to this question later. Again, for the moment, we will simply take relative m-complexity to be related simply to the number of morphemes of a particular utterance. With these narrow definitions, m- and p-complexity appear to be independent, but the interaction between the two will become apparent later.

This is a competing motivations explanation, since it claims that the pressures that these factors bring to bear on the selection of relative clauses are opposed. Consider the following Malagasy examples (from Keenan 1972b — N.B. these illustrative examples are quoted directly from this source):

(12) ny vehivavy izay nividy ny vary ho an' ny ankizy
the woman REL bought the rice for the children
'the woman who bought the rice for the children'

(13) a. * ny vary izay nividy ho an' ny ankizy ny vehivavy
the rice REL bought for the children the woman
'the rice which the woman bought for the children'

b. ny vary izay novidin' ny vehivavy ho an' ny ankizy
the rice REL bought+PASS the woman for the children
'the rice which the woman bought for the children'

(14) a. * ny ankizy izay nividy ny vary (ho an) ny vehivavy
the children REL bought the rice (for) the woman
'the children who the woman bought the rice for'

b. ny ankizy izay nividianan' ny vehivavy ny vary
the children REL bought+CIRC the woman the rice
'the children who the woman bought the rice for'

(12) is an example of a subject relative in Malagasy. (13a) shows that object relativisation in Malagasy is ungrammatical. This raises the question of how speakers get round the problem of presenting the message in (13a) without using an ungrammatical relative. The solution in Malagasy is to promote the object to subject using a passive and then relativising on the derived subject (13b). This structure is morphologically marked with respect to the non-passivised equivalent since it involves extra passive morphology on the verb, hence it has a higher m-complexity. Similarly, Malagasy oblique relatives (14a) are ungrammatical (as we should expect given the AH). Instead, speakers can use another promotion-to-subject construction (14b). Here, a "circumstantial" affix is attached to the verb that promotes the oblique object to subject. Again, this clearly involves an increase in m-complexity.

Here, then, is a case where avoidance of some relative causes an *increase* in m-complexity, but a *decrease* in p-complexity.[3] Thus, the two complexity motivations are in competition. For hearers the prime concern is a low p-complexity. For speakers, the prime concern is a low m-complexity (in line with a principle of least effort). If these two complexity measures can be in conflict, then we can say that speakers and hearers are in conflict. Notice, however, that even *if* speakers were altruistic, the conflict between the need to reduce parsing complexity and the need to keep utterances short would remain "within" speaker as opposed to between the speaker and the hearer. In other words, for a competing motivations explanation, the crucial factor is that there be conflicting measures of functional pressure, wherever these pressures make their influence felt in the cycle of production and parsing.

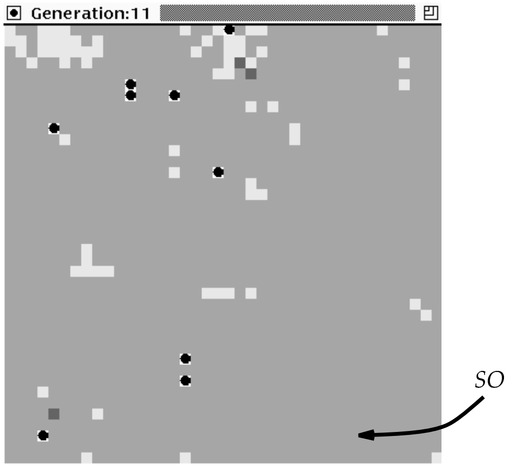



Figure 2: An example run of the simulation with high p-complexity, showing predominance of $SO$.

# 6 Testing the competing motivations

In order to see what effect m-complexity has on the simulation, the way in which I-language is mapped onto utterances (transformation **T1**) needs to be adjusted. In other words, it is too simplistic to say that speakers produce utterances in line with their I-language states; instead, the probability of producing morphologically simpler forms should be weighted higher than the higher m-complexity variants. To do this, a variable $c_m$ is introduced that represents the relative m-complexity of $S'$ or $O'$ with respect to $S$ or $O$. Reflecting the fact that speakers prefer relatives, and hearers disprefer relatives, especially object relatives, (given our characterisation of m- and p-complexity) the parameters of the simulation are:

(15)

| variable | values | interpretation |
|---|---|---|
| $c_m$ | $c_m > .5$ | m-complexity of RC variants |
| $c_S$ | $c_S > .5$ | p-complexity of subject RC |
| $c_O$ | $c_O > c_S > .5$ | p-complexity of object RC |

Depending on the initial conditions, one of two results emerge depending on the relative magnitude of m-complexity and p-complexity. If m-complexity is high, then the end result is languages of type $S'O'$ only, whereas if p-complexity is high, the end result is languages of type $SO$ only (see figure 2). Obviously, with neither starting condition does the hierarchy emerge.

Although this result seems to suggest that the competing motivations hypothesis has failed, this rests on an incorrect assumption about the nature of the values of the variables in (15). These variables are set to certain values at the start of the simulation and remain the same for all points in the simulation space and over time. However, it is not plausible to say that the relative magnitude of m-complexity and p-complexity will be invariant for languages. To see why, compare the Malagasy examples above with some Malay examples also from Keenan (1972b):

(16) Ali bunoh ayam itu dengan pisau
Ali killed chicken the with knife
'Ali killed the chicken with the knife'

(17) a. * pisau yang Ali bunoh ayam itu dengan
knife REL Ali killed chicken the with
'the knife that Ali killed the chicken with'

b. pisau yang Ali gunaka untok membuno ayam itu
'the knife that Ali used to kill the chicken'

Malay is unable to relativise on obliques (16-17a), however there is no way in which to promote the oblique to subject to reduce the p-complexity of oblique relatives, as in Malagasy (13b). When Keenan's informants were asked to produce an equivalent to the English oblique relative, they gave a paraphrase such as (17b). The relative cost to the speaker of such a paraphrase compared to a promoted equivalent is obvious in terms of the number of morphemes (i.e. effort) involved in its production.

As well as paraphrase and promotion, circumlocution is another strategy for avoiding relatives. Consider variants to the English (19a) and (18a).

(18) a. I watch the batsman who England selected.

b. I watch the batsman who was selected by England.

(19) a. I watch the team which Hick plays cricket for.

b. *I watch the team which was played cricket for by Hick.

c. I watch this team — Hick plays cricket for them.

(18b) is the promoted variant of (18a), but the passive is not available to promote the oblique and reduce p-complexity (19b). Another option in this case is to use something like (19c) which does not have a relative at all.

These examples make it clear that the relative m-complexity of variants is language specific. In certain languages like Malagasy, there is a morphologically simple voice system that enables easy promotion to subject (i.e. promotion introduces little extra morphological material into the utterance). Malay, on the other hand, has a less well developed system, and cannot promote obliques. English can promote some NPs, but the passive is morphologically more complex than in Malagasy. To sum up, then, the relative magnitude of m- and p-complexity is not universally fixed, rather it is affected by the systems made available by the rest of the language and may vary over time.

In order to model this, the simulation was adjusted so that every few generations the relative magnitude of m- and p-complexity was adjusted for a random language type. This involved introducing another parameter that expressed the probability of a change occurring each generation, but the value of this parameter does not seem to be critical. The result of this seemingly small change in the simulation is profound. Instead of settling down to a static end state like the other simulation runs, the state of the simulation 'world' is constantly changing. Large groups are formed, as in Levin's simulation, but at the boundaries of these groups something akin to borrowing occurs, and language types move across space, and change in prominence over time. A few of the generations in a typical run of the simulation are shown in figure 3. The most important feature of the results is that all language types are well represented *except for* $S'O$. This type is in the figure as black, and takes up about one quarter of the initial space, by generation 10, however, there is almost none of this type displayed. Over a long run, the other three types (indicated for the final generation) share the space roughly between them. The implicational universal has emerged.

To summarise, the results from the three simulation experiments are:

**p-complexity only:** static end state – $S'O'$

**p- and m-complexity, fixed:** static end state – $S'O'$ *or* $SO$

**p- and m-complexity, variable:** dynamic state – $S'O'$ $SO'$ *and* $SO$

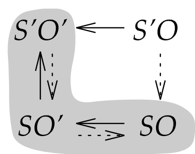
These results lend strong support to a competing motivations analysis within a selection model where the magnitude of the selection pressures is variable. The next section discusses how this result can be generalised to the rest of the AH, and gives an explanation for the subject relative universal (4).

# 7 Dynamic typology

In order to understand what the simulation is doing, we need a theory of how dynamic processes give rise to universal constraints. In other words, if we understand what types of changes are likely to occur when the simulation is in one state, then is there a way to calculate what universals will emerge? Borrowing from Greenberg (1978) we will use *type graphs* in order to answer this question.

A type graph is a graph whose nodes are states in a language typology, and whose arcs are possible transitions between those states. So, for the example discussed above, there will be four nodes in the type graph: $S'O'$, $S'O$, $SO'$ and $SO$. As we have seen, which transitions between these states are possible depends on the relative magnitude of m- and p-complexity. This is represented by two different types of arc: solid ones for when p-complexity considerations are paramount, and dotted ones for when m-complexity is greater:

(20)




If we follow the transitions on this graph we can see what happens to a language in the simulation given a particular initial state. So, if a language relativises on subjects and objects, and the m-complexity of RC variants is low, then the next state of the language will be subject-only relativisation, and then neither subject nor object relativisation.[4] Considering only the solid arcs on the graph, then the situation is equivalent to the first run of the simulation where m-complexity was not considered. It is clear that the inevitable end state will be $S'O'$ since once a language is in this state, then it cannot escape. This is termed a *sink* by Greenberg (1978:68). Similarly, if only the dotted arcs are considered, then $SO$ is a sink. This explains why the second simulation run always ended up at one of these two end states depending on the initial conditions. One can imagine that a kind of tug-of-war is going on between the two types of complexity. If one team in a tug-of-war is consistently better than

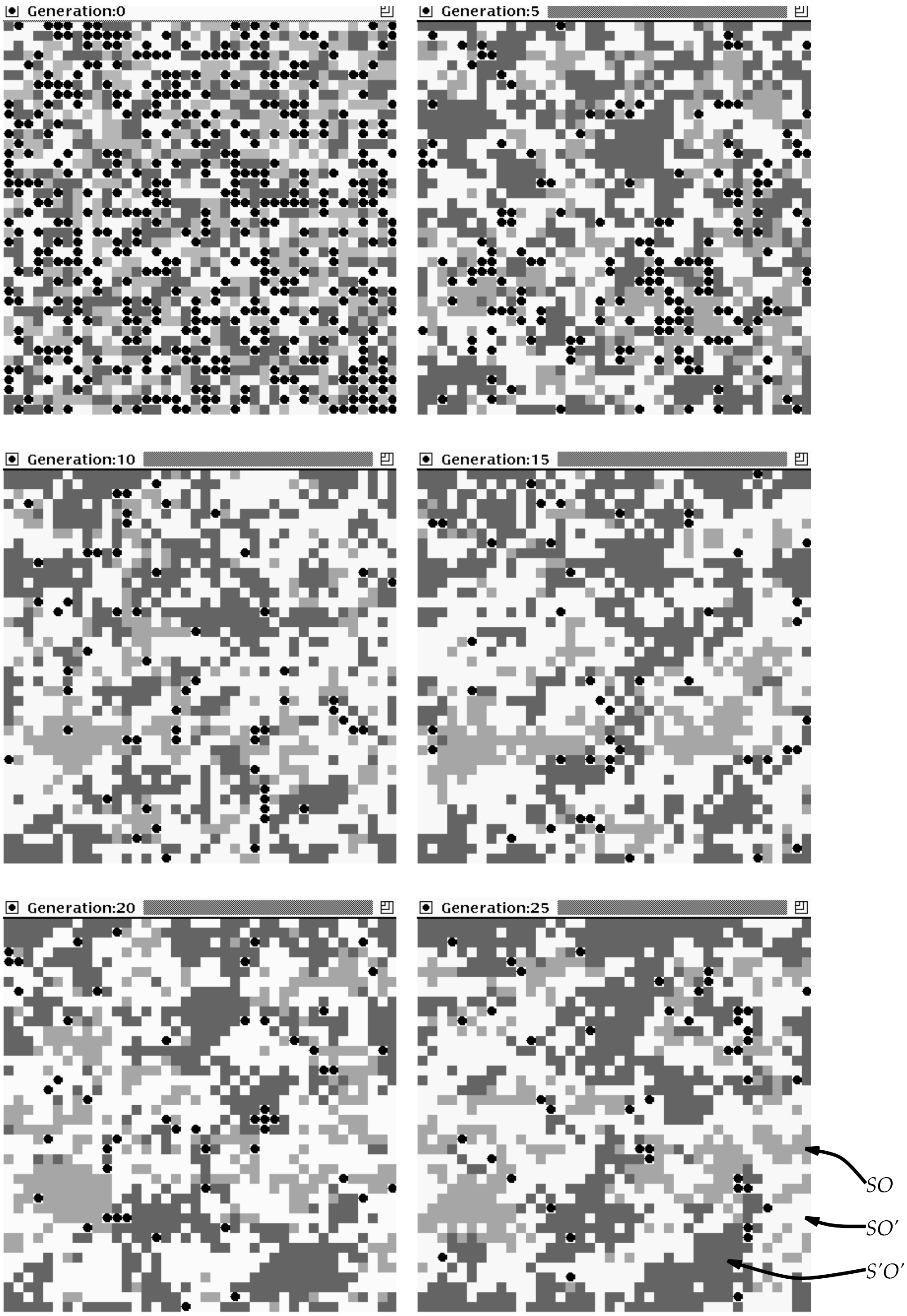



Figure 3: An example run of the simulation with shifting complexities. Note that number of the $S'O$ type (here highlighted as black circles) is reduced rapidly from the initial condition. (Proportion of $S'O$ is 27% at generation 0, and 3% at generation 25.)

another, then that team will always win — the rope will always be pulled to one end. Over time, the influence of the other team on the *end result* of the game is irrelevant.

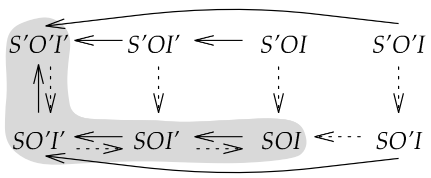
If both types of arcs are considered, then the implicational universal emerges: languages end up in the shaded region of the graph. An informal definition of areas of type graphs that correspond to universals is given below:

(21) The language types that are predicted to occur are the set of nodes that belong to strongly connected sub-graphs whose members are only connected to other members of the sub-graph.

A node $a$ is 'connected' to $b$ if there is an arc from $a$ to $b$, or if there is an arc from $a$ to $c$ and $c$ is connected to $b$. A graph is 'strongly connected' if for every node $a$ and every node $b$ in the graph $a$ is connected to $b$ (and vice versa). So, in (20) all the nodes in the shaded region are connected to each other, but once languages are in this region they cannot escape from it.

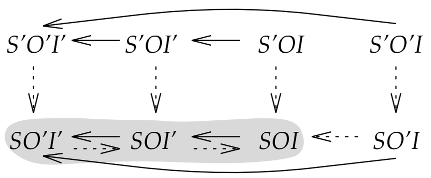
The graph can be extended to other positions on the hierarchy. So, for example, (22) is the graph for the first three positions on the AH: subject, direct object and indirect object. Again, the universal that is predicted by the definition above is shaded:

(22)




The shaded regions in the graphs above are indeed what the accessibility hierarchy predicts; however, they do not correspond to what is found in reality. This is because of the separate subject relative universal (4) which states that all languages relativise on subjects. This is a case where the type graph theory can be used to look for a possible explanation. The smallest change that can be made to the graph above to bring it in line with the observed universal is to remove the solid arc leading from $SO'I'$ to $S'O'I'$ (i.e. remove the hearer-driven change that makes subject relatives ungrammatical):

(23)




In fact, it seems that this might indeed be the correct modification to the previous explanation. Recall that languages typically provide a number of possible ways of 'avoiding' a particular relative clause construction. One of the least morphologically complex of these strategies is the promotion-to-subject strategy exemplified by the Malagasy examples (16–17). This strategy is *not* available to avoid subject relatives, however, and even if the language allowed demotion this would not be a viable option since it would increase the p-complexity of the relative clause. So, this calls into question an idealisation in the design of the simulation: namely, that relative m- and p-complexity shifts randomly. If promotion is unavailable for subjects, then the average relative m-complexity of constructions that avoid subject relativisation will be higher than for other positions. Selection by the speaker – in terms of m-complexity – will thus be more likely for this position.

## 8 Case coding and complexity

So far only *primary* relativisation strategies have been considered. These are strategies for relativisation that are used for subjects according to Keenan and Comrie's definition. However, languages often make use of different strategies for relativisation on lower positions on the hierarchy. It turns out that the competing motivations approach makes some interesting predictions for the distribution of these strategies.

Two broad types of relativisation strategy are examined in early work:

(24) **The case coding taxonomy:** (adapted from Comrie & Keenan 1979 and Keenan & Comrie 1977) A strategy for relativisation is case-coding (or [+case]) if a nominal element is present in the restricting clause which case marks the relativised NP at least as explicitly as is normally done in simple sentences.

An example of a [–case] strategy in Arabic relativisation is given by Keenan & Comrie (1979:333):

(25) al- rrajul ya'raf al- sayeda allati nayma
the man knows the woman REL sleeps
'The man knows the woman who is sleeping'

Here the relative marker does not code for the case of the NP being relativised, and there is no extra nominal element with the clause that marks its case. Object relativisation in Arabic is [+case], however (Keenan & Comrie 1979:333):

(26) al- walad ya'raf al- rajul allathi darabat -hu al- sayeda
the boy knows the man that hit him the woman
'The boy knows the man whom the woman hit'

In this example, the case is coded by the resumptive pronoun *-hu* within the restrictive clause. Another example of a [+case] strategy is given by standard written English direct object relativisation:

(27) The boy knows the man whom we saw.

Here, the relative pronoun marks the relativised NP as a direct object. Notice that the commonly used relative markers (*who, which, that*) occurring in subject and direct object relativisation can all be used for both those positions and are thus [–case], since they do not explicitly code the case of the relativised NP.

In these examples, and universally, [+case] strategies occur lower on the AH than [–case] strategies. This is predicted by the theory outlined in this paper if we include a notion of information content in the definition of p-complexity. When defining the p-complexity of RCs it was argued that complexity must be relative to a particular psycholinguistic operation – namely the association of the trace, or resumptive pronoun, with the head noun.

The complexity of this association task may be ameliorated by providing (typically redundant) information relating to the grammatical function of the embedded element. Hawkins (1994:45–46) supports a similar analysis: the 'conservation of logical structure' hypothesis of Keenan (1972a). This states that resumptive pronouns make the correspondence between surface structures and logical-semantic structures of relative clauses more transparent, and therefore make processing easier. However, this analysis only covers resumptive pronouns, whereas a treatment in terms of redundancy of information covers the full range of possible [+case] strategies.

The two types of strategy differ with respect to both m- and p-complexity:

(28) a. **[+case]** High relative m-complexity (extra nominal element increases morphological markedness), low relative p-complexity.

b. **[–case]** Low relative m-complexity, high relative p-complexity.

At first blush, this seems to make no predictions about the distribution of strategies. Again, m- and p-complexity are in conflict. However, the relative markedness of the two strategies changes down the accessibility hierarchy:

(29) a. **Change in relative m-complexity:** The typical m-complexity of an RC high on the hierarchy will be lower than that of one low on the hierarchy, therefore any increase of m-complexity will be more marked high on the hierarchy.

b. **Change in relative p-complexity:** The low positions on the hierarchy have higher p-complexity, so it is less likely that a form that increases p-complexity further will survive to the trigger on these positions.[5]

It is apparent that case-coding represents a trade-off between an increase in m-complexity and a decrease in p-complexity. For positions low on the hierarchy the balance is in favour of selection in terms of p-complexity (hearer selection) giving [+case] strategies, whereas positions high on the hierarchy favour selection in terms of m-complexity (speaker selection) giving [–case] strategies.

## 9 Beyond [+/–case]

Tallerman (1990) revises the definition of [+case] to include examples where the relativised NP is marked without an explicit nominal element. The motivation for this is to analyse examples of consonantal mutation in Welsh – which disambiguate the function of the relativised NP – as [+case]. The new definition also includes strategies that explicitly mark the grammatical function of the relativised NP by word order (e.g. English):

(30) **Case coding strategies:** (Adapted from Tallerman 1990, 293) A strategy for relativisation is case-coding or [+case] if it explicitly signals the grammatical function of the relativised NP. (Not necessarily with a nominal element.)

In fact, this means that most languages use solely [+case] strategies, in Tallerman's sense, unless word order produces ambiguous relative clauses. Welsh provides examples where there are both [+case] and [–case] strategies, since the basic word order is VSO (Tallerman 1990:296).

(31) y bachgen a welodd *t* y ci *t*
the boy COMP saw-3SG the dog
'the boy who saw the dog' or
'the boy who the dog saw'

In this example, the *t*s mark the possible positions for the trace, yielding the two possible readings respectively. This, then, is [–case] relativisation. As mentioned above, Welsh consonantal mutation provides a [+case] strategy (Tallerman 1990:300):

(32) y bachgen a welodd *t* **g**i
the boy COMP saw-3SG dog(+MUT)
'the boy who saw a dog'

(33) y bachgen a welodd **c**i *t*
the boy COMP saw-3SG dog(-MUT)
'the boy who a dog saw'

Put simply, there is a morphophonemic set of changes in Welsh known as soft mutation which occurs on some segments in certain environments, including directly following a noun phrase. Wh-traces are included in the set of triggering environments, hence the mutation of the initial segment in *ci* above.

An interesting feature of Tallerman's definition of [+case] is that it allows us to go beyond the simple case-coding strategies with opposition between speaker and hearer and look in more detail at the interaction of m-complexity and cross-linguistic distribution. Firstly, a further definition:

(34) **Zero-morpheme strategy:** A strategy that is case-coding (in Tallerman's sense) but uses no extra morphemes ('nominal elements') for case-coding is a zero-morpheme strategy.

Hence, Welsh soft mutation is a zero-morpheme strategy. Since zero-morpheme strategies are case-coding, with low relative p-complexity, but *without* the concomitant increase in relative m-complexity, we can predict that zero-morpheme strategies will be used as high on the accessibility hierarchy as they can be.[6] This is indeed true in the Welsh case. If the so-called word order strategies in the sample of Maxwell 1979 are taken into account, then this is further support for this prediction since they are all primary strategies.

We can extend the prediction about zero-morpheme strategies by formulating a hierarchy of strategies that is ranked in terms of m-complexity:

(35) **Strategy hierarchy:** [+case] strategies may be ordered with respect to the typical relative m-complexity of case-coding, such that a complex or 'weighty' strategy occurs low on the hierarchy:

Zero-morph > Case-coding Relative Pronoun >? Anaphoric Pronoun[7]
(> Clitic Doubling etc.)

The lower the strategy is on this hierarchy, the lower on the accessibility hierarchy that strategy will occur cross-linguistically.

This hierarchy is rather speculative since there has been no typological research that categorises strategies to this level of detail. The study of Maxwell (1979) refines the Keenan/Comrie sample by categorising strategies as word-order, relative-pronoun and anaphoric-pronoun, among others. Maxwell's categorisation is obviously not motivated by morphological complexity and we must be cautious of any support that his work provides. However, it is interesting to note that the distribution of anaphoric pronoun strategies in the sample is skewed significantly lower on the accessibility hierarchy than that of the relative pronoun strategies.[8]

Even within one language, we can find support for the strategy hierarchy. Looking again at Welsh, Tallerman (1990:313) notes that a pronominal strategy can be used for some direct objects, some non-direct objects and genitives. A clitic doubling strategy, however, is only available for some non-direct objects and genitives. This distribution is expected since the clitic doubling strategy (37) has a higher m-complexity that simple retention of an anaphoric pronoun (36):

| (36) | y | bachgen | y | gwnaeth | y | ci | ei | weld |
|---|---|---|---|---|---|---|---|---|
| | the | boy | COMP | did-3SG | the | dog | 3MSG | see |

'the boy that the dog saw' (Tallerman 1990:302)

| (37) | y | papur | roeddwn | i'n | edrych | arno | fo |
|---|---|---|---|---|---|---|---|
| | the | paper | COMP-was-1SG | I-PROG | look | at-3MSG | it(3MSG) |

'the paper that I was looking at' (Tallerman 1990:306)

## 10 Conclusion

This paper has attempted to show how a close look at exactly how functional pressures give rise to universals is an important step in assessing the adequacy of functional explanations. In particular, the explanation of hierarchical (implicational) universals by appealing to a gradient processing hierarchy is inadequate. Instead, a competing motivations explanation is required.

The solution to the problem of linkage (1) is a *selection model* of linguistic adaptation – a type of invisible hand process. The power of the selection model lies in its ability to explain features of language distribution that *appear* to be designed for ease of communication, without imputing any 'teleology of purpose' in the users of the language – a charge levelled at some functional explanations (see for example Lass 1980 for a critical review). So, although it looks as if speakers have selected relativisation possibilities in such a way as to aid the comprehension of utterances, this is an illusion. The ordered nature of the hierarchy is simply an inevitable emergent property of the dynamics of language use, acquisition and propagation. The simulation described in section 4 is a useful tool to test the predictions of the model, and provides a simple demonstration of emergence.

Clearly, the explanation for the accessibility hierarchy put forward in this paper makes no assumption of speaker altruism. I believe this is the correct null hypothesis. If speakers *did* take hearers' needs into account in the choice between two competing forms, then the 'black box' that maps intended messages onto forms would need to be more complex. Instead of making choices moment to moment on the basis of some kind of principle of least effort, calculations would need to be made about the effort of parsing the final structure. Notice, however, that the model is not incompatible with the assumption of speaker altruism, simply that it is not required in order to explain the accessibility hierarchy.

With respect to the accessibility hierarchy specifically the competing motivations hypothesis and the selection model allow us to go further in examining the hierarchy itself. Using type graphs to analyse what the selection model is doing reveals a plausible explanation for the subject relative universal without requiring any additional assumptions. Looking in more detail at the metrics of complexity involved in selection suggests a theoretically motivated categorisation of relativisation strategies in terms of morphemic weight. Deciding exactly how this weight should be measured depends on an understanding of the psychological processes involved in the production of utterances. In order to test the predictions in section 9 more precisely, we need to know whether to count morphs or morphemes, case and $\phi$-features or all features in a paradigm. Answers to these questions are out of the scope of this paper, but a fruitful area of inquiry may be in the literature on morphological processing and specific language impairment (e.g. Clahsen 1989; Lenormand *et al.* 1993).

Finally, I believe that the model shows that functional explanations for language universals are compatible with generative approaches to language acquisition. The discussion on linguistic selection explicitly allows the representation medium of linguistic knowledge to be autonomous from language function. This does *not* rule out functional explanations of constraints on cross-linguistic variation. Indeed, in some circumstances the functional approach and the 'formal' approach to universals may both be required in order to explain cross-linguistic patterns (Kirby 1996).

## Acknowledgements

I would like to thank Jim Hurford, Diane Nelson, Louise Kelly, Ronnie Cann and Jack Hawkins for their helpful comments during the preparation of this paper. The opinions expressed, and any errors made, are however my own.

## Notes

1. In fact, some sisters may be excluded from the calculation if the language has flatter configurational structure. In this case, morphological case contributes to the calculation of structural complexity. For example, in languages without VPs, nominative marked NPs may be included as sisters of an accusative, but not vice versa. See Hawkins 1994, 27–28, for discussion.

2. In some situations, the simulation in this paper acts in the same way, since there is only selection between competing variants when there is sufficient variability in the environment.

3. The relative clauses are subject relatives, and thus have smaller structural domains. Hawkins (1994:31) explicitly states that the calculation of structural complexity should relate to the position of the co-indexed element inside the clause "in its original (d-structure) position" in an attempt to provide a unified account of promotion-type relatives such as (13b) and non-promoted relatives. However, there are reasons why we should be wary of this approach and, at least as a first approximation, use a definition that refers to the surface position. Firstly, arguments about the d-structures of the constructions are theory-specific, and Hawkins does not state allegiance to a particular polystratal syntactic model. Secondly, and more importantly, one of the results of Keenan & Hawkins (1987) work is that when errors are made repeating relatives, then the errors tend to be towards relatives on lower positions on the hierarchy. The majority of errors made repeating relativised direct objects were RCs on the subject of a passive; the majority of errors made repeating relativised subjects of passives, however, where RCs on direct objects. A possible explanation, is that the former case is a response to p-complexity (the RC was mis-parsed), whereas the latter is a response to m-complexity (a simpler paraphrase is produced).

4. This graph only shows what will happen all things being equal – in other words, if there is sufficient random variation in the environment to allow speakers and hearers to freely select variant forms. The simulation described in the last section does *not* make this assumption, however, since variation is drawn from other languages which are also following paths through the type graph.

5. Notice that the asymmetry between speaker and hearer selection here is explicable given that speakers make selection 'choices' by comparing the two variants directly, whereas hearers/acquirers do not have direct access to a comparison of the two forms at the point of selection.

6. This will generally mean that they will be used for subject relativisation (i.e. they will be primary strategies), however it is conceivable that a zero-morpheme strategy may be constrained in other ways so that it cannot be freely selected for on every position on the hierarchy.

7. The ordering of these two strategies may depend on an assessment of the degree to which the two types of pronoun encode $\phi$-features across languages.

8. A Mann-Whitney U test gives us a significance level of $p < 0.005$, but this level may partially be due to the sampling technique.

## References


BYBEE, JOAN. 1988. The diachronic dimension in explanation. In *Explaining Language Universals*, ed. by John A. Hawkins. Blackwell.

CHOMSKY, NOAM. 1986. *Knowledge of Language*. Praeger.

CLAHSEN, H. 1989. The grammatical characterisation of developmental dysphasia. *Linguistics* 27.897–920.

COMRIE, BERNARD, & EDWARD KEENAN. 1979. Noun phrase accessibility revisited. *Language* 55.649–664.

CROFT, WILLIAM. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

——. 1993. Functional-typological theory in its historical and intellectual context. *Sprachtypologie und Universalienforschung* 46.21–22.

DUBOIS, JOHN. 1987. The discourse basis of ergativity. *Language* 64.805–855.

GIVÓN, TALMY. 1979. *On Understanding Grammar*. Academic Press.

GREENBERG, JOSEPH. 1978. Diachrony, synchrony, and language universals. In *Universals of Human Language 1, Method and Theory*, ed. by Joseph Greenberg, 61–91. Stanford.

HALL, CHRISTOPHER. 1988. Integrating diachronic and processing principles in explaining the suffixing preference. In *Explaining Language Universals*, ed. by John A. Hawkins. Blackwell.

——. 1992. *Morphology and Mind: a Unified Approach to Explanation in Linguistics*. Routledge.

HAWKINS, JOHN A. 1988. Explaining language universals. In *Explaining Language Universals*, ed. by John A. Hawkins. Basil Blackwell.

—— 1992. Innateness and function in language universals. In *The Evolution of Human Languages*, ed. by J.A. Hawkins & M. Gell-Mann. Addison-Wesley.

—— 1994. *A performance theory of order and constituency*. Cambridge University Press.

HURFORD, JAMES. 1990. Nativist and functional explanations in language acquisition. In *Logical Issues in Language Acquisition*, ed. by I.M. Roca, 85–136. Foris Publications.

KEENAN, EDWARD. 1972a. The logical status of deep structures. In *Proceedings of the Eleventh International Congress of Linguists*, ed. by L. Heilmann. Societa editrice il Mulino.

——. 1972b. Relative clause formation in Malagasy. In *The Chicago Which Hunt*, ed. by Paul Peranteau, Judith Levi, & Gloria Phares, 169–189. Chicago Linguistics Society.

——, & BERNARD COMRIE. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.63–99.

——, & BERNARD COMRIE. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55.333–351.

——, & SARAH HAWKINS. 1987. The psychological validity of the accessibility hierarchy. In *Universal Grammar: 15 essays*, ed. by Edward Keenan, 60–85. London: Croom Helm.

KELLER, RUDI. 1994. *On language change: the invisible hand in language*. Routledge.

KIRBY, SIMON. 1994a. Adaptive explanations for language universals: a model of Hawkins' performance theory. *Sprachtypologie und Universalienforschung* 47.186–210.

——. 1994b. Constraints on constraints, or the limits of functional adaptation. In *Proceedings of the Edinburgh Linguistics Department Conference '94*, ed. by Simon Kirby, 48–54.

——. 1996. *Function, Selection and Innateness: the Emergence of Language Universals*. PhD thesis, University of Edinburgh.

LASS, ROGER. 1980. *On Explaining Language Change*. Cambridge University Press.

LENORMAND, M.T., L.B. LEONARD, & K.K. MCGREGOR. 1993. A cross-linguistic study of article use by children with specific language impairment. *European Journal of Disorders of Communication* 28.153–163.

LEWONTIN, R. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press.

MACWHINNEY, B., & C. PLEH. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition* 29.95–141.

MAXWELL, DANIEL. 1979. Strategies of relativization and NP accessibility. *Language* 55.352–371.

NEWMEYER, FREDERICK J. 1994. Competing motivations and synchronic analysis. *Sprachtypologie und Universalienforschung* 47.67–77.

SOBER, ELLIOT. 1984. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. MIT Press.

TALLERMAN, MAGGIE. 1990. Relativisation strategies: NP accessibility in Welsh. *Journal of Linguistics* 26.291–314.