Logical Issues in Language Acquisition

I.M. Roca (ed.)

1990
FORIS PUBLICATIONS
Dordrecht - Holland/Providence RI - U.S.A.

Nativist and Functional Explanations in Language Acquisition

James R. Hurford
University of Edinburgh

1. PRELIMINARIES

1.1. Setting and Purpose

Current theories of language acquisition and of linguistic universals tend to be polarised, adopting strong positions along dimensions such as the following: formal (or nativist) versus functional; internal versus external explanation; acquisition of language versus acquisition of communication skills; specific faculté de langage versus general cognitive capacity.

As with many enduring intellectual debates, there is much that is convincing and plausible to be said on each side. Some works are very polemical, apparently conceding little merit in the opposing point of view. Some so-called 'functional' explanations of language universals, which appeal to properties of performance mechanisms, e.g. the human parser, miss the important point that these mechanisms are themselves innate and as much in need of explanation as the properties of the linguistic system. Another class of proposed functional explanations for language universals, which appeal to the grammaticalisation of discourse patterns, fail to locate this mechanism in the life-cycle of individual language-knowers. On the other hand, some nativist explanations imply that they are complete, having finally wrapped up the business of explaining language acquisition, missing the point that the demand for explanations never ceases, and that the 'solution' to any given puzzle immediately becomes the next puzzle.

The appearance of a direct confrontation between nativist and functional styles of, or emphases in, explanations of language acquisition and linguistic universals was greater in the 1970s than it has been recently, as Mallinson (1987) emphasises. Golinkoff and Gordon (1983) give a witty, but fairly accurate, historical account of the pendulum-swings and emphasis-shifts in the debate since the inception of generative grammar. Regrettably, the embattled spirit of the barricades survives in some quarters, as in Newmeyer's (1983) review of Givon (1979), itself a sharp polemic, and in the exchange between Coopmans (1984) and Hawkins (1985).

In an area where polemic is so rife, the truth-seeker can be distracted or misled by a number of false trails which it is as well to be able to

recognize in advance. The following are some types of distraction to be ever vigilant for: (1) Unannounced theory-laden use of everyday terms, such as 'language', or 'universal' (for instance, using 'language' to mean just the unmarked core grammar, or excluding phonology); (2) The assumption of a monolithic research enterprise, such that a criticism of any single aspect of it is taken as a blanket attack on the whole; and (3) Sheer mistaking of an opposing position, taking it to be something other (even the opposite) of what it really is (a distressingly frequent type of mistaking involves elementary failure to distinguish between 'all' and 'some' in an opponent's exposition).

I assume that the readership of this book will not consist wholly, or even largely, of convinced generative linguists, but will include people such as psychologists studying language acquisition, linguists with a more anthropological emphasis, philosophers who ponder issues of language structure and use, sociolinguists, and theorists of historical language change, to all of whose work logical issues in language acquisition are relevant. Being concerned with outlining a synthesis of approaches accessible to workers in these different areas, my points will typically be at a quite general level, and I will often resort to quoting relevant work from the various fields. The distinctions I discuss will tend to be broad distinctions between domains of study, rather than the finer distinctions identified by workers within domains. Seekers after very specific proposals about models and mechanisms will not find them here. But, at this general level, I will propose a model for the interaction of language use and language acquisition, in which I believe all students of language, from psycholinguists through 'core' linguists to sociolinguists and historical linguists, will be able to identify a part which is theirs.

A colleague has likened this attempt at synthesis to waving a flag in the no-man's-land between two entrenched armies shooting at each other, with the consequent likelihood of finding oneself full of bullet-holes. But the military metaphor is, one hopes, inappropriate to scholarly work. Synthesizing, integrating work must be attempted. This is not to discourage any individual researcher from trying to mount a strong case that such-and-such an aspect of language should be attributed to the influence of the innately structured LAD (or alternatively to what I shall call the Arena of Use), nor to dissuade any rival researcher from trying to demolish such a case, on theoretical or empirical grounds. Indeed such efforts, locally partisan as they are, are the sine qua non of the growth of knowledge in the field. What I am trying to discourage is a dismissive globally partisan, academically totalitarian, kind of view, that holds that explanations from innateness (or, for the opposing partisan, from use) are simply not worth serious consideration, on either theoretical or empirical grounds.

1.2. Glossogenetic and Phylogenetic mechanisms

The dimension of diachrony, only skimpily treated in previous discussions, provides a coherent background within which function and innateness can be consistently accommodated. Functional explanations of language acquisition can be compatible with nativist explanations, provided one gets the timescale right. The much-debated dichotomy, innate versus functional, is a red herring. The basic dichotomy is, rather, phylogeny versus ontogeny, and also the related nature versus culture. Function is not 'opposed' to any elements in these dyads, but exerts its influence on all.

The issue of the relation between linguistic development and other (cognitive, social, etc.) experience can be set in different timescales, short-term or long-term. Such experiences may be directly involved with linguistic development within the time-span of an individual's acquisition of his language, a period of a few years; or, at the other extreme, the outcomes of experiences of members of the species over an evolutionary timescale lead to the natural selection of individuals innately equipped to acquire systems with particular formal properties. The idea of short-term (ontogenetic or glossogenetic) timescales versus long-term (phylogenetic) timescales in explanations for linguistic facts is important to an overall view of the relation between function and innateness. The term 'glossogenetic' reflects a focus on the development and history of individual particular languages; language-histories are the rough cumulation, over many generations, of the experiences of individual language acquirers. The biological endowments of successive generations of language acquirers in the history of a language do not differ significantly, and so linguistic ontogeny, and its cumulation, language history, or glossogeny, are to be distinguished from linguistic phylogeny, the chronologically vastly longer domain, in which biological change, affecting the innate language faculty, takes place. After the present section of preliminaries, the second and main section of this paper will be devoted to the short-term, onto- or glossogenetic mechanism of functional influence on language form.

A detailed exposition of the phylogenetic mechanism of functional influence on language form is, unfortunately, too long to be included in this collection of papers, and is to be published elsewhere (Hurford, 1991). The phylogenetic mechanism is mentioned briefly by Chomsky and Lasnik (1977:437), but although their note has been echoed by various subsequent authors (e.g. Lightfoot, 1983:32, Newmeyer, 1983:113, Lasnik, 1981:14, Wexler, 1981:40), it has not initiated an appropriate strand of research into functional explanations of language universals at the level of evolution of the species.

Despite acceptance of the premise that functional explanations for linguistic universals do operate at the level of evolution of the species,

remarkably little further gets done about it. Contributions from linguists, of whatever theoretical persuasion, (e.g. Lightfoot's section "Evolution of Grammars in the Species" (Lightfoot, 1983:165-169) and Givon's chapter "Language and Phylogeny" (Givon, 1979:271-309)) remain sketchy, superficial, and anecdotal.

On the other hand, a more promising sign is Pinker and Bloom's (1990) paper, in which they systematically address some of the major skeptical positions (e.g. of Piattelli-Palmarini, 1989, Chomsky, and Gould) concerning natural selection and the evolution of the language faculty. Several other articles (Hurford, 1989, 1991a, 1991b; Newmeyer, forthcoming) make a start on working out proposals about how quite specific properties of the human language faculty could have emerged through natural selection.

To whet the reader's appetite, without, I hope, appearing too enigmatic or provocative at this stage, I give here a short paragraph with a diagram (Figure 1), sketching the phylogenetic mechanism, and a table (Table 1), summarising the major differences between the glossogenetic and the phylogenetic mechanisms. Deep aspects of the form of language are not likely to be readily identifiable with obvious specific uses, and one cannot suppose that it will be possible to attribute them directly to the recurring short-term needs of successive generations in a community. Here, nativist explanations for aspects of the form of language, appealing to an innate LAD, seem appropriate. But use or function can also be appealed to on the evolutionary timescale, to attempt to explain the structure of the LAD itself.

The phylogenetic explanatory scheme I envisage is as follows:

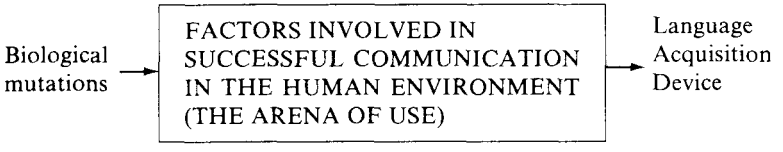

Fig. 1.

Here biological mutations plus functional considerations constitute the *explanans*, and the LAD itself constitutes the *explanandum*. The LAD is part of the species' heredity, the result of mutations over a long period.

TWO TYPES OF FUNCTIONAL EXPLANATION		
	GLOSSOGENETIC (Sec.2 of this paper)	PHYLOGENETIC (Hurford, 1989, 1991b)
Usefulness felt:	In short term (every generation)	In long term (evolutionary timespan)
Transmission:	Cultural	Genetic
Knowledge determined by data:	Typically, well determined	Typically, poorly determined
Innovation by:	Invention, creativity of individuals	Biological mutation
Typical explanandum:	Language-specific	Universal
Competition in Arena of Use:	Between languages (Ln vs Ln+1)	Between classes of languages
Motivating analogy:	Language as a TOOL	Language as an ORGAN

Table 1.

Much of the present paper will be an extended commentary on the rubrics in this table, especially those in the 'Glossogenetic' column. Before getting down to the details of the glossogenetic mechanism in Section 2,there are a couple of general preliminary plots to be staked out, in the remainder of this section.

1.3. Competence/performance, I-Language/E-language

Explanations differ according to what is being explainɗd. This is a truism. But much discussion of 'explaining linguistic phenomeŋa' uses that phrase to smother an important distinction, the distinction between grammaticality and acceptability (competence and performance, I-language and E-language). The distinction is central to the Chomskyan enterprise, and has been a frequent target of attack, or source of misgivings. In the literature, for instance, one finds widely-read authors writing:

> "The distinction between competence and performance - or grammar and speaker's behavior - is ... untenable, counterproductive, and nonexplanatory". (Givon, 1979:26)

> "The borderline between the purely linguistic and the psychological aspects of language ... may not exist at all". (Clark and Haviland, 1974:91)

> "There is a whole range of different objections from sociolinguists, sometimes querying the legitimacy of drawing the [competence/performance] distinction at all". (Milroy, 1985:1)

Givon's book is still widely discussed, Herb Clark is an influential psychologist, and Lesley Milroy speaks for a body of sociolinguists for whom the competence/performance distinction itself is still a current issue. In the context of their expressed doubts about competence/performance (alternatively I-language/E-language), and concomitantly grammaticality/acceptability, it is relevant to reassert this distinction. Despite such doubts and attacks, I will maintain here that many clear cases of the distinction exist, while conceding that there are borderline linguistic phenomena whose classification as facts of grammar or facts of use is at present problematic.

Some early, perhaps overhasty, conclusions claiming to have explained aspects of grammar in functional terms can now be reinterpreted as explaining phenomena more peripheral to the grammatical system, such as stylistic preference, or acceptability. For instance, this is how Newmeyer (1980:223-226) depicts Kuno's various functional explanations: 'Kuno's approach to discourse-based phenomena has gradually moved from a syntactic one to one in which the generalisations are to be stated outside of formal grammar' (Newmeyer, 1980:224). Such reinterpretation follows shifting (and, one hopes, advancing) theories of the boundary between grammatical phenomena proper and acceptability and style.

For concreteness, I will give some examples, all for Standard English, of how I assume some relevant phenomena line up:

(1) GRAMMATICAL, BUT OF PROBLEMATIC ACCEPTABILITY
Colourless green ideas sleep furiously.
The mouse the cat the dog chased caught ate some cheese.
The horse raced past the barn fell.

(2) UNGRAMMATICAL, AND OF PROBLEMATIC ACCEPTABILITY
*He left is surprising.
*The man was here is my friend

(3) UNGRAMMATICAL, BUT OFTEN ACCEPTABLE
*He volunteered three students to approach the Chairman
*She has disappeared the evidence from her office

A degree of relative agreement between individuals, and certainty within individuals, about the above examples does not mean that there can't be genuine borderline cases. There may well be slight differences between individuals in their genetically inherited language faculties[1], and the input data is certainly very variable from one individual to another, as is the wider social context of language acquisition. And (any individual's instantiation of) the language acquisition device itself may not be structured in such a way as to produce a classification of all possible wordstrings with respect to their grammaticality.

This classification of patterns of linguistic facts as grammatical or otherwise does not depend, circularly, on the kind of explanatory mechanism one can postulate for them, but rather primarily in practice (though by no means wholly in principle) on that classical resource of generative grammar, native speaker intuitions of grammaticality (themselves not always easily accessible).

In fact, from a linguist's viewpoint, the sentences (1-3) constitute a heterogeneous bunch, conflating much more interesting distinctions which these very sentences, if aptly exploited, could well emphasise, for instance, grammaticality versus parsability, grammaticality versus first-choice parsing strategies, semantically correct versus conceptually empty sentences etc. But I am stressing here a more basic point. The grammaticality/acceptability distinction, paralleling the competence/performance (I-language/E-language) distinction, is an absolutely crucial foundation upon which the further much more interesting distinctions can be elaborated. Only if it is accepted can one progress to the more interesting distinctions. In this paper my concern is to investigate the relationships obtaining between the domain of grammar, on the one hand, and nongrammatical, e.g. processing-psychological and social, domains, on the other hand. For my purposes, as it turns out, these other domains can, at a broad general level, be lumped together, so far as their role in potential functional explanations for aspects of linguistic competence is concerned, although obviously a study with a different focus of attention would immediately separate and distinguish them. Sociolinguistics, pragmatics and discourse analysis, and psycholinguistics are disciplines with highly divergent goals and methodological styles. (Thus 'functional explanation' is likely to be interpreted in different ways by sociolinguists and psycholinguists.) Chomsky is entirely right in emphasising that a language (E-language) is an artifact resulting from the interplay of many factors. Where I differ from his judgement is in my belief that this artifact is of great interest, that it is susceptible to systematic study (once its diverse component factors are identified), and that it can in fact affect grammatical competence (I-language).

Given the grammaticality/acceptability distinction, and a classification,

however tentative, of linguistic facts according to this distinction, the search for explanations must provide appropriate explanatory mechanisms for the different kinds of linguistic phenomena. The explanatory task for grammaticality facts can be couched fairly naturally in terms of language acquisition: 'How does a person acquire a particular set of intuitive judgements about wordstrings?' But the explanatory tasks for the various diverse classes of acceptability facts are not naturally couched in terms of language acquisition.

Different kinds of questions require different kinds of answers, but this does not mean that, for example, perceptual strategies can ultimately play no part in explaining how a child acquires certain grammaticality judgements. And, conversely, it does not mean that grammatical facts (competence) can play no part in processing. (To the linguist convinced of the grammaticality/acceptability distinction, processing necessarily involves grammatical facts.) But as the mechanisms which give rise to competence obviously differ in their 'end products' from the mechanisms which give rise to acceptability facts (performance), the details of the two kinds of mechanisms themselves must be different. The reasons for distinguishing competence from performance are very well set out, in partly Saussurean terminology, by Du Bois (1985).

> "Saussure (1959:11-23, 191ff) demarcates sharply between what he calls internal linguistics, the study of *langue*, and external linguistics, which encompasses such significant fields of study as articulatory phonetics, ethnographic linguistics, sociolinguistics, geographical linguistics and the study of utterances (discourse?), all of which deal with positive facts.
>
> Classical structuralism thus establishes a gulf between the two spheres, so that structuring forces or organizing principles which operate in the one domain will not affect the other. Though this formulation will be seen to be too one-sided, given its assumption that *langue* is in principle independent of structuring forces originating outside it, I will suggest that the distinction between internal linguistics and external linguistics nevertheless remains useful and in fact necessary. I will draw on this distinction to show how certain phenomena can be at the same time unmotivated from the generative synchronic point of view and motivated from a genuinely metagrammatical viewpoint which treats grammars as *adaptive systems*, i.e. both partially autonomous (hence systems) and partially responsive to system-external pressures (hence adaptive). This will be fruitful only if we recognise the existence of *competing motivations*, and further develop a theoretical framework for describing and analysing their interaction within specified contexts, and ultimately for predicting the resolution of their competition. This (panchronic) approach to metagrammar is part of the developing theory of what has been called the ecology of grammar (Du Bois, 1980:273)." (1985:343-344).

The ecological metaphor is also taken up, independently, in Hurford (1987). While I am in sympathy with Du Bois's approach, and regard it as an admirably clear statement of the system/use dilemma that modern linguistics has forged for itself, I believe Du Bois has not gone as far as

he might in considering the *ontology of grammar*. That is, he still tends, in a Saussurean way, to treat grammatical systems as abstractions, with their own laws and principles, without locating them in the minds of speakers. And he does not locate the mechanism of grammaticisation in the Chomskyan LAD, which, I believe, is where it belongs.

Sociolinguists' difficulties with the competence/performance distinction stem largely, according to Milroy, from the problem of language variation. And several current models of language acquisition respond to the pervasive fact of variation by proposing that the linguistic competence acquired is itself variable. Thus Macken (1987) proposes that acquired grammars are partly 'algebraic' and partly 'stochastic'. And the 'competition model' (Bates and MacWhinney, 1987; MacWhinney, 1987a,b) assumes that:

> "... the 'steady state' reached by adults also contains patterns of statistical variation in the use of grammatical structures that cannot be captured by discrete rules". (Bates and MacWhinney, 1987:158)

This echoes early attempts to reconcile sociolinguistic variation with generative grammar's view of competence; cf. Labov's (1969) idea of 'variable rules', its development by Cedergren and Sankoff (1974), and critical discussion by Romaine (1982:247-251).

The facts of linguistic variation and gradual linguistic change lead Kroch (1989) to propose another possibility, distinct from both the 'single discrete competence' and the 'probabilistic competence' views.

> "If we ask ourselves why the various contexts of a linguistic alternation should, as a general rule, be constrained to change in lock step, the only apparent answer consistent with the facts of the matter is that speakers learning a language in the course of a gradual change learn two sets of well-formedness principles for certain grammatical subsystems and that over historic time pressures associated with usage (presumably processing or discourse function based) drive out one of the alternatives". (Kroch, 1989:349)

This echoes a long tradition in linguistics (cf. Fries and Pike, 1949).

It is hard, perhaps impossible, to distinguish empirically between a situation where a speaker knows two grammars or subsystems, corresponding, say, to 'New Variety' and 'Old Variety', and a situation where a speaker knows a single grammar or subsystem providing for a number of options, where these options are associated with use-related labels, 'Old' and 'New'. Plural competences would certainly be methodologically more intractable to investigate, presenting a whole new, and more difficult, ball-game for learnability theory, for instance. On the other hand, plural competences do presumably arise in genuine cases of bilingualism, and so the LAD is equipped to cope with internalizing more than one grammar

at a time. Perhaps plural competences are indeed the rule for the majority of mankind, and the typical generative study of singular monolithic competence is a product of concentrating on standardised languages (a point made by Milroy). The question is forced on us by the pervasive facts of statistical patterning in sociolinguistic variation, even in the usage of single individuals, and language change. And the question is highly relevant to language acquisition studies, as McCawley (1984:435) points out: 'Do children possess only one grammar at a time? Or may they possess multiple grammars, corresponding to either overlapping developmental stages, or multiple styles and registers?'

In what follows I will simply assume that statistical facts belong to the domain of performance and pragmatics (e.g. rules of stylistic preference or, more globally, rules of 'code choice'), whereas facts of acquired adult grammatical competence are not to be stated probabilistically. I do not claim to have *argued* this assumption, or demonstrated that the variation problem must be handled in this way. But one cannot explore all the possibilities in one article, and I shall explore here how the interplay of grammar and use might be envisaged, if one banishes probabilities from the realm of competence. The research challenge then appears as the twin questions: 'How does all-or-nothing competence give rise to phenomena in which statistical distributions are apparent?' and 'How does exposure to variable data result in all-or-nothing competence?' Possibly, these are the wrong research questions to ask, but the only way to find out is by seeing how fruitful theorising along these lines turns out to be. Other researchers may pursue other assumptions in parallel. In a later subsection (2.3), I will discuss the phenomenon of grammaticalisation, in which, over time, a statistical pattern of use (as I assume it to be) gets fixed into a nonstatistical fact of grammar.

1.4. The ambiguity of 'functional'

Opponents of nativist explanations for linguistic universals often contrast the Chomskyan doctrine of an innate Language Acquisition Device with a form of explanation labelled 'functionalist'. Such functionalist explanations point to the *use* of language as accounting for the properties of linguistic systems. But typically in such accounts, one of two distinct aspects of 'use' is emphasised. Hyman identifies this ambiguity clearly:

> "Unfortunately, there is disagreement on the meaning of 'functional' as applied in this context. While everyone would agree that explanations in terms of communication and the nature of discourse are functional, it became evident in different presentations at this workshop that explanations in terms of cognition, the nature of the brain, etc., are considered functional by some but not by other linguists. The distinction appears

> to be that cognitive or psycholinguistic explanations involve formal operations that the human mind can vs. cannot accommodate or 'likes' vs. 'does not like', etc., while pragmatic or sociolinguistic explanations involve (formal?) operations that a human society or individual within a society can vs. cannot accommodate or likes vs. does not like". (Hyman, 1984:67-8)

The same kind of distinction between types of functional explanation is noted, but labelled differently, by Bever (1975):

> "There have been two major kinds of attempts to explain linguistic structure as the result of speech functions. One I shall call the 'behavioural context' approach, the other the 'interactionist' approach. The 'behavioural context approach' argues that linguistic patterns exist because of general properties of the way language is used and general properties of the mind. The interactionist approach argues that particular mental mechanisms guide and form certain aspects of linguistic structure". (Bever, 1975:585-6)

And Atkinson (1982) makes approximately the same distinction between alternative reductive explanations for language acquisition, which he labels 'cognitive reductions' and 'social reductions'.

The distinction between cognitive and social reductions (Atkinson's terms), between explanations based on an interactionist approach and those based on a behavioral context approach (Bever's terms) is by no means clear-cut. All humans have cognition and all engage in social relations; but social relations are experienced and managed via cognition (and perception). Social relations not thus mediated by perception and cognition are hard, if not impossible, to conceive. A good illustration of a 'social' principle with substantive 'cognitive' content is the Gricean Maxim of Manner, 'Be perspicuous'. This maxim is generally (by now even conventionally!) held up as an example of the influence of social considerations on language use. But 'Be perspicuous' clearly has psychological content. What is perspicuous to one kind of organism may be opaque to an organism with different cognitive structuring. As Grice's work is widely known, this statement in terms of a Gricean maxim is adequate to make the point of the interpenetration of cognitive and social 'functional' factors. Sperber and Wilson's (1986) Relevance Theory, which claims to have supplanted the Gricean model with a deeper, more general, more explanatory theory of social communication through language, lays great stress on the individual psychological factor of processing effort.[2] Speakers' discourse strategies are jointly motivated by what hearers find easy to understand (a cognitive consideration) and by a desire to communicate efficiently (a social consideration). Functional explanations can indeed have the different emphases which Hyman, Bever, and Atkinson all identify, but cognitive and social factors are often intermingled and not easy to separate.

An explanation of some aspect of language structure is functional to the extent that it provides an account relating that aspect of structure to some purpose for which language is used, or to some characteristic of the users or manner of use facilitating achievement of that purpose. The canonical form of a functional explanation is as in (4).

(4) X has form F because X is used by U and/or for purpose P.

where some clear connection between F (the putatively useful form) and U (the user) and/or P (the purpose) is articulated. The connection between form and user or purpose need not be immediate or direct but may be mediated in some way, provided the plausibility of the connection is not thereby lost. As a simple concrete example, consider a spade. Parts of its form, e.g. the sharp metal blade, relate directly to the intended purpose, digging into the earth, but other aspects of its form, e.g. its handle and its manageable weight, relate more directly to the given (human) characteristics of the user. Separating out which aspects of spade-design are purpose-motivated and which user-motivated is not easy; likewise it can also be difficult to separate out social (purpose-motivated) functional explanations of language form from psychological (user-motivated) functional explanations.

For the purpose of exploring the relationship between nativist and functional explanations of linguistic phenomena, it will in fact be convenient to continue to deal in terms of a single functional domain, which has both cognitive and social components. This domain, which I will label the 'Arena of Use' and discuss in the next section, is contrasted with the 'internal' domain, the domain of facts of grammar. The Arena of Language Use must figure in any explanation of language form that can reasonably be called a 'functional' explanation.

2. GLOSSOGENETIC MECHANISM OF FUNCTIONAL INFLUENCE ON LANGUAGE FORM

2.1. The Arena of Use

The familiar nativist scheme for explaining the form of grammatical knowledge is shown in *Figure 2.*

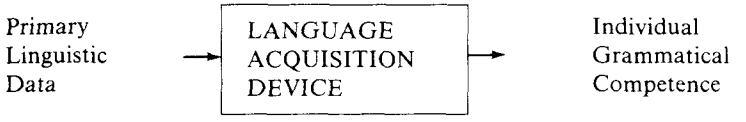

Fig. 2.

In this scheme, the grammatical competence acquired by every individual who learns a language conforms to a pattern determined by innate psychological properties of the acquirer. These innate characteristics are influential enough to impose significant patterning, not obviously discernible in the primary linguistic data, on the acquirer's internalized grammar. Whatever the primary linguistic data (within the range normally experienced by young humans) the competence acquired on exposure to it conforms to the specifications built into the Language Acquisition Device. So, across languages and cultures, adult language-knowers carry what they know in significantly similar forms, studied under the heading of Universal Grammar (UG).

The short-term functional mechanism by which nongrammatical factors can in principle contribute to linguistic phenomena, and ultimately to grammatical competence, can be represented by an extra component added to the Chomskyan diagram (Figure 2), as in Figure 3 below.

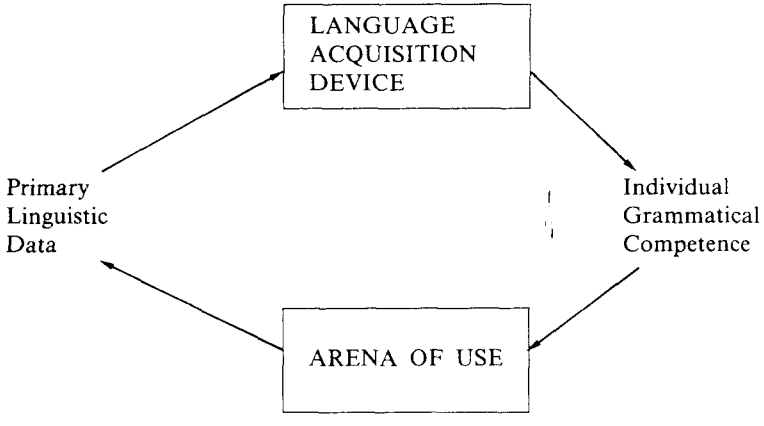

Fig. 3.

What is the Arena of Use? Well, it is non-grammatical, that is it contains no facts of grammar, although it relates to them. And some of it is non-

psychological, in the sense of being outside the domain of individual mental processes, although it receives input from these, and provides material for them. The Arena of Use does have some psychological ingredients, including those directly involved in linguistic performance. The Arena of Use is where utterances (not sentences) exist. The Arena of Use is a generalisation for theoretical purposes of all the possible non-grammatical aspects, physical, psychological, and social, of human linguistic interactions. Any particular set of temporal, spatial, performance-psychological and social coordinates for a human linguistic encounter is a point in the Arena of Use. So, for example, an address or point in the Arena of Use that I happen just to have visited might be approximately described by the phrase: 'Jim Hurford, sitting in his living room at noon on January 6th, having some cognitive trouble composing an elegant written sentence (strictly an inscription) about the Arena of Use, for an unknown readership, assumed to consist of assorted academic linguists, sociolinguists and psycholinguists.' Another address in the Arena of Use might be 'Mrs Bloggs, at the greengrocer's, asking loudly, since the grocer is a bit deaf, for 2lbs of leeks'. The Arena of Use is where communication takes place. It embraces human relationships, the ways in which we organise our social lives, the objects that it is important to us to communicate about, the kinds of message it is important for us to transmit and receive. Other creatures, built differently from ourselves, would conduct their communication in, and have it shaped by, a different (though probably partly similar) Arena of Use. So, note, the Arena of Use is itself partly, in fact very largely, a product of our heredity (part of our 'extended phenotype', in Dawkins' (1982) phrase).

The Arena of Use, like UG, has both absolute and statistical properties. A full description of the Arena would specify a definite, obviously infinite, range of possibilities, the coordinates of possible communicative interactions between people using language; and, within this range, the likelihood of the various possibilities being realised would be projected by various principles, in a way analogous to the role played by a theory of markedness within UG. Obviously, we are no nearer to a full description of the Arena of Use than we are to a full description of UG, but central aspects of the nature of the Arena are nevertheless relatively easily accessible for hypothesis and consideration.

The Arena of Use is emphatically *not* 'everything there is (provided it has no grammatical import)'. And it is certainly not equivalent to, or coextensive with, 'the world' or 'the environment'. I will try to clarify. In the first place, the world is 'out there', existing somehow outside our perceptions. (I assume this, being a Realist and not an Idealist.) In knowing the world, we impose categories on it that are to a great extent our own constructs, though they presumably mesh in some way with the ways things

really are 'out there'. The Arena of Use is not populated by just whatever exists out there, but (in part) by entities that exist-as-some-category. The relevant idea is put thus by Edie (commenting, as it happens, on Husserl):

> "For Husserl no 'object' is conceivable except as the correlate of an act of consciousness. An 'object' is thus never a thing-in-the-world, but is rather something apprehended about a thing; objects are things *as intended, as meant, as* taken by a subject". (Edie, 1976:5, his emphasis, quoted by Fraser, 1989:79)

> And Fraser elaborates this eloquently:
> "Many of the Objects that we encounter are presented to us *as what they are* through a filter of our language and culture, rather than being constituted anew by each Subject on the basis of individual experience". (Fraser, 1989:121)

The inclusion in the Arena of Use of abstract objects constituted through language illustrates how it itself (or better, its instantiation in a particular historic language community) is something dynamic and developing. Traugott (1989) discusses a diachronic tendency for meanings of words to develop from concrete denotations of objects and states of affairs to more abstract denotations 'licensed by the function of language' (Traugott, 1989:35). Thus 'everything there is', 'the world', or 'the environment' is quite different from the Life-worlds of individual subjects, speakers of a language; the Life-worlds are in some ways richer, in some ways poorer, than the actual world, although clearly there is a degree of correspondence. The Arena of Use includes the sum of entities (and classes of entities) in the Life-worlds of individual Subjects (speakers) that these subjects can talk about. (This excludes strictly private psychological entities that might be quite real for many individuals, but which they cannot talk about. One reason for not being able to talk about some experience is the lack of appropriate words and/or grammatical constructions, which is why creative writers sometimes resort to novel forms of expression.)

The Arena of Use is not just a union of sets of (classes of) entities. It has structure and texture (much of which remains to be articulated by pragmatic theory). Some (but not all) of its structure is statistical, deriving from the salience (or otherwise) for numbers of speakers of particular classes of entities. Prominent classes of entities in the Arena of Use are those that everyone talks about relatively frequently. Other aspects of the Arena of Use are what Fraser, following Husserl and Heidegger, calls 'points of view' and what Lakoff (1986:49) calls 'motivation'. Humans have purposes, and employ language to manipulate other speakers to help them to achieve those purposes. There are ways in which this is typically done, which gives rise to the taxonomies of Speech Act theory. In fact, any theory of pragmatics contributes to a theory of the Arena of Use, and the categories postulated by pragmatic theorists, such as speaker,

hearer, overhearer, deixis of various types, utterance, situation of utterance, illocution, perlocution, implicature, etc., etc. are all theoretical categories forming part of our (current) picture of the Arena of Use. The Arena of Use is in part the subject matter of pragmatics, and it would clearly be wrong to say that it is 'just everything there is', 'the world', or 'the environment'. If this were so, nothing would distinguish pragmatics from, say, a branch of physics.

As for the usefulness of coining the expression 'Arena of Use', my purpose is to focus attention on a vital link in the transmission of language from one generation to the next. Chomsky's similarly ambitious expression 'Language Acquisition Device' has played an enormously important role in focussing theorists' attention on the other important link in the cycle. Clearly, it would have been unimaginative and counterproductive several decades ago to dismiss that expression on the grounds that it simply meant 'child'.

It should be clear that the role of the Arena of Use is *complementary to* that of the LAD, not, of course, in any sense proposed as an alternative to it. And, in fact, just because of this complementarity, studies of UG actually *need* systematic information about the Arena of Use. Thus Lightfoot (1989a:326) is forced to resort to a 'hunch' about whether a particular hypothetical social scenario is plausible or 'too exotic', when conducting an argument about whether "The existence of N′ might be derived from a property of UG or ... might be triggered by the scenario just sketched". (Grimshaw, 1989:340, complains about the lack of independent evidence backing such hunches.) Obviously it would be too much to expect a theory of the Arena of Use to give a direct answer to this specific question, but, equally obviously, the more systematic a picture of the Arena of Use we can build up, the less we will need to rely on hunches about what the input data available to the child may be. For instance, observation of actual caretaker behaviour is a necessary empirical support to the axiom of 'no negative evidence' central to UG and learnability theory (Lightfoot, 1989a:323-324, Grimshaw and Pinker, 1989:341-342, inter alios but cf. Saleemi, this volume). And a pragmatic theory of *why* caretakers give little or no negative evidence, if we could get such a theory, would neatly complement the UG and learnability theories.

The products of an individual's linguistic competence are *filtered* by the Arena of Use. In the Chomskyan scheme, the LAD acts partly as a filter. The child in some sense disregards the properties of utterances in the Primary Linguistic Data that do not conform to his innate (unconscious) expectations, the characteristics that cannot be interpreted in terms of the structure already possessed (a function recently emphasised and elaborated on by Lightfoot, 1989a. Similarly, the Arena of Use acts as a filter. Not all the products of an individual's competence serve any

useful purpose, and these are either simply not uttered, or uttered and not taken up by interlocutors.

At the level of discourse, the filtering function of the Arena is accepted as uncontroversial. A coherent discourse (monologue or dialogue) is not just any sequence of sentences generated by a generative grammar. The uses to which sentences are put when uttered determine the order in which they may be strung together. With the usual reservations about performance errors, interruptions, etc., sequences which do not serve useful purposes in discourse do not occur in the Primary Linguistic Data to which the child is exposed.

At the level of vocabulary, the filtering function of the Arena is also uncontroversial. Words whose usefulness diminishes are uttered less frequently, eventually falling out of use. When they fall out of use, they are no longer present in the PLD and cannot pass into the competences of new language acquirers. What words pass through the cycle in Figure 3, assuming their linguistic properties present no acquisition difficulties, is almost entirely determined by considerations of use. I grant that the relation between vocabulary and use is far from simple, as academic folktales about Eskimo words for snow (cf. Pullum, 1989, Martin, 1986), and Arabic words for camel might lead the gullible to believe. But there is a large body of scholarship, under the various titles of ethnographic semantics, ethnoscience, and cognitive anthropology (cf Brown, 1984, for a recent example), building up a picture of the relation between the structure of a community's vocabulary and its external environment. Clearly the usefulness of words is one part of this picture. One example from Brown is:

> "The fact that warm hues cluster with white and cool hues with dark contributes to the likelihood that languages will make a "macro-white"/"macro-black" distinction in the initial encoding of basic color categories. A utilitarian factor may also contribute to this development. Basic color categories become important when people develop a need to refer to colors in a general manner. An initial "macro-white"/"macro-black" contrast is highly apt and useful since it permits people to refer to virtually all colors through use of general terms". (Brown, 1984:125)

At the level of a semantico/pragmatic typology of sentences, it also seems plausible that the existence of universal types is perpetuated through the mediation of the Arena of Use, rather than of the LAD. The three-way distinction declarative/interrogative/imperative reflects the three most salient types of speech act used in human interaction. This taxonomy and its grammatical realisation is probably passed on to successive generations via ample exemplification in the Arena of Use, necessitating no extraordinary innate powers of extrapolation from skimpy data by the LAD. A theory of the acquisition of grammatical competence, such as UG, makes

available a range of syntactic forms. Without reference to pragmatics, which provides a classification of the uses to which sentences may be put, there is no account of why *three* (as opposed to five or nineteen) types of syntactic structure are salient and typically assigned to different uses. What UG cannot account for, without recourse to a pragmatic theory is this:

> "There is a wealth of cross-language evidence showing the existence of three or four syntactic structures which code prototypical speech acts in any language:
>
> (a) Declarative
> (b) Imperative
> (c) Interrogative
> (i) WH-question
> (ii) Yes/No question.
>
> It is hard to find a language in which some "norm" does not exist for (a), (b), (ci) and (cii), i.e. some structural-syntactic means for keeping these four prototypes apart." (Givon, 1986:94)

We can think of UG as providing a theory of the formal/structural resources, or space, available to humans for the expression of useful distinctions. Obviously, a theory of just what distinctions are useful (pragmatic theory, theory of the Arena of Use) is also needed. That is, "One must then strive to discover the underlying *socio-psychological* parameters which define the multi-dimensional space within which speech-act prototypes cluster". (Givon,1986:98) Then, interesting discussion can proceed on how specific features of use tend to select specific structural features of sentence form for their expression. Givon's suggestion is that there is an iconic relation between the syntactic forms and their functions, but this clearly needs more fleshing out. Downes (1977) is an interesting paper suggesting why the imperative construction, in particular, occupies the area of syntactic space that it does, e.g. with base form of the verb and suppressed subject. A theory of grammar, such as UG, can make available sentences with null subjects and with base verb forms, but the question arises: Why are these sentences, in particular, typically used to get people to do one's bidding?. My intention is not to dispense with the theory of UG. But the allocation of individual aspects of a phenomenon to a theory of grammar-acquisition or a theory of use must be considered on its merits. Perhaps the assignment of 2nd person to the null subject of imperatives, for example, is a blank that UG can afford to leave to a theory of use. This is in fact what Beukema and Coopmans suggest:

> "... the position is occupied by a case-marked empty element associated with an empty topic, which receives the interpretation of addressee from the discourse". (Beukema and Coopmans, 1989:435)

Beside the declarative/interrogative/imperative pragmatic typology, one could also cite the categories of person and number, which recur in all grammars, as motivated by factors in the Arena of Use. Hawkins puts it concisely:

> "Innateness is not the only factor to which one can appeal when explaining universals. Certain linguistic properties may have a communicative/functional motivation. If every grammar contains pronouns distinguishing at least three persons and two numbers (cf. Greenberg 1966:96), then an explanation involving the referential distinctions that speakers of all languages regularly need to draw is, *a priori*, highly plausible". (Hawkins, 1985:583)

The facts of grammatical person are not quite so simple. Foley (1986:66-74) (while subscribing to the same functional explanation as Hawkins for distinctions of grammatical person) mentions languages without 3rd person pronouns, and Mühlhäusler and Harré (1990) claim that even 1st versus 2nd person, as usually understood, is not universal. Nevertheless Hawkins' point stands; it is not surprising that 'the referential distinctions that speakers of all languages regularly need to draw' cannot be described by a simple list, but rather require description in statistical terms of significant tendencies.

Hawkins gives a number of further plausible examples, which I will not take the space to repeat. In a more recent, and important, contribution the same author accounts for universal tendencies to grammaticalise certain word orders in terms of certain (innate) parsing principles:

> "The parser has shaped the grammars of the world's languages, with the result that actual grammaticality distinctions, and not just acceptability intuitions, performance frequencies and psycholinguistic experimental results, are ultimately explained by it. This does not entail, however, that the parser must also be assumed to have influenced innate grammatical knowledge, at the level of the evolution of the species, as in the discussion of Chomsky and Lasnik (1977). Rather, I would argue that human beings are equipped with innate *processing* mechanisms in addition to innate *grammatical* knowledge, that the grammars of particular languages are shaped by the former as well as by the latter, and that the cross-linguistic regularities of word order that we have seen in this paper are a particularly striking reflection of such innate mechanisms for processing. The evolution of these word order regularities could have come about through the process of language change (or language acquisition): the most frequent orderings in performance, responding to principles such as EIC [Early Immediate Constituents, a parsing principle], will gradually become fixed by the grammar. One can see the kinds of grammaticalization principles at work here in the interplay between "free" word order and fixed word order within and across languages today. The rules or principles that are fixed by a grammar in response to the parser must then be learned by successive generations of speakers". (Hawkins, 1990:258)

Another case of the influence of phenomena in the Arena of Use on patterns of grammar is discussed in detail by Du Bois (1987). This study attributes the existence of ergative/absolutive grammatical patterning to preferences in discourse structure. The study has the merit of providing substantial statistics on these discourse preferences. The link between such discourse preferences and ergative grammatical patterning is argued for very plausibly. And Du Bois answers the obvious question 'Why are not all languages ergative?' by appealing (again plausibly, I believe) to competing motivations, discourse pressures in several directions.

As a final example here of the contribution of the Arena of Use to the form of linguistic phenomena, I cite certain properties of numeral systems, in particular the universal property of being organised on a base number (often 10). There is no evidence that children somehow innately prefer numeral expressions organised in the familiar way using as a base-word the highest-valued available numeral word in the lexicon. Rather, the modern streamlined systems have evolved over long historical periods because of their practical usefulness, and they have to be deliberately inculcated into children. (This argument concerning numerals is pursued in detail in Hurford, 1987, where a computer simulation of the social interactions leading to the emergence of the base-oriented structure of numeral systems is presented.)

In summary, the Arena of Use is the domain in which *socially useful* and *cognitively usable* expressions are selected to fit the worldly purposes of hearers and speakers. The Arena contributes to the form of languages in a way complementary to the contribution of the Language Acquisition Device. Languages are artifacts resulting from the interplay of many factors. One such factor is the LAD, another is the Arena of Use. The aspects of languages accounted for by these two factors are complementary. As a first approximation, one might guess that the aspects of languages due to the LAD are relatively deep, or abstract, whereas the aspects due to the Arena of Use are relatively superficial, in the sense in which the terms 'deep' and 'superficial' are typically used by generative grammarians. The terms 'deep' and 'superficial' tend to be rhetorically loaded, and imply triviality for superficial aspects of language. One need not accept such a value judgement. The deep characteristics of languages most convincingly attributed to the Language Acquisition Device are those to which the 'poverty of stimulus' argument applies, that is, characteristics which are not likely to be encountered in a sampling of primary linguistic data. Such deep characteristics are thus those which are actually *least* characteristic of languages, in any normal pretheoretical sense, in the sense of being least obvious. Thus the theoretical style typifying research into the contribution of the Arena of Use is to be expected, in the first place at least,

to be more 'superficial' than research into UG and the LAD. But the intrinsic interest of such a theory is not thereby diminished.

A full and helpful discussion of the uses of 'deep' by generative grammarians and others, and of the misunderstandings which have arisen over the term, is to be found in Chapter 8 of Chomsky (1979). Putting aside the use of 'deep' as a possible technical term applied to a level of structure (which I am *not* talking about here), the term 'deep' can be applied either to theories and analyses or to phenomena and data considered pretheoretically. Those aspects of languages due to the LAD seem, at first pretheoretical blush, to be 'deep', to require theories of notable complexity to account for them. These aspects of a language's structure are subtle; they are not the most obvious facts about it, and, for instance, probably get no attention in courses teaching the language, even at an advanced level. Exactly this point is stated by Chomsky:

> "We cannot expect that the phenomena that are easily and commonly observed will prove to be of much significance in determining the nature of the operative principles. Quite often, the study of exotic phenomena that are difficult to discover and identify is much more revealing, as is true in the sciences generally. This is particularly likely when our inquiry is guided by the considerations of Plato's problem, which directs our attention precisely to facts that are known on the basis of meager and unspecific evidence, these being the facts that are likely to provide the greatest insight concerning the principles of UG". (Chomsky, 1986:149)

This subtlety in acquired knowledge after exposure to data in which the subtlety is not obviously present accounts for the rise of complex theories of language acquisition. On the other hand, those aspects of languages due to the Arena of Use (many of which would be located in the periphery of grammars by generativists, like irregular and suppletive morphological forms) seem not to require anything so complex – they are much less underdetermined by data, and thus require no invocation of special deep principles to account for their acquisition.

My reservation about not necessarily accepting the value judgements implicit in much current usage of 'deep' stems from the association that has now become established between 'deep' and the language-acquisition problem. In a theory of language cast as a theory of language acquisition, or 'guided by the considerations of Plato's problem', the term 'deep' is applied, naturally, to aspects of language whose acquisition apparently necessitates deep analyses. In this sense, the question of how children acquire irregular morphological forms, for example, is relatively trivial, not deep; the child just observes each such irregularity individually and copies it. (Well, let's say for the sake of argument that the right answer really is as simple as that, which it isn't, clearly.) That's not a deep answer, so the question, apparently, wasn't deep. But seen from another perspective,

the same aspects of language could well necessitate quite deep analyses. If one casts a theory of language as a theory of communication systems[3] operating within human societies (systems transmitted from one generation to the next), then the problem of acquisition is not the only problem one faces. The kind of question one asks is, for instance: Why do these communication systems (languages) have irregular morphological forms?, Why do languages have words for certain classes of experience, but not for others? And the answer to these questions may be quite deep, or at least deeper than the answers to the corresponding acquisition questions. (A similar argument is advanced in Ch.1 of Hurford, 1987)

Figure 3, introducing the Arena of Use, is actually a version of a diagram given by H. Andersen (1973). Andersen's diagram looks like this:

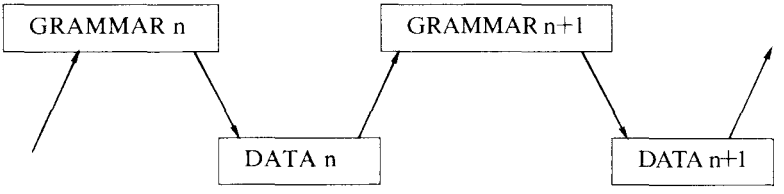

Fig. 4.

Andersen is interested in the mechanisms of linguistic change, and makes the basic point that grammars do not beget grammars. Grammars give rise to linguistic data, which are in turn taken and used as the basis for the acquisition of grammars by succeeding generations. Lightfoot (1979) argues on these grounds that there can be no theory of linguistic change expressed as a theory *directly* relating one grammar to a successor grammar. A theory attempting to predict the rise of new grammars from old grammars purely on grounds internal to the grammars themselves would be attempting to make the spurious direct 'horizontal' link between GRAMMAR *n* and GRAMMAR *n*+1 in Figure 4.

The zigzag in Figure 4 could be extended indefinitely across the page, representing the continuous cycle through acquired grammars and the data they generate. The LAD belongs on the upward arrows between data and grammars. The Arena of Use belongs on the downward arrows, between grammars and data. In fact Figures 3 and 4 both represent exactly the same diachronic spiral, merely differing in emphasis. Figure 3 is simply Figure 4 rotated and viewed 'from one end'.

Pateman (1985), also drawing on this work of Andersen's, expresses very neatly the relationship I have in mind between grammars and social or cultural facts:

> "... through time the content of mentally represented grammars, which are not in my view social objects, comes to contain a content which was in origin clearly social or cultural in character". (Pateman, 1985:51)

George Miller also expresses the same thought concisely and persuasively:

> "Probably no further organic evolution would have been required for Cro-Magnon man to learn a modern language. But social evolution supplements the biological gift of language. The vocabulary of any language is a repository for all those categories and relations that previous generations deemed worthy of terminological recognition, a cultural heritage of common sense passed on from each generation to the next and slowly enriched from accumulated experience". (Miller, 1981:33)

It is worth asking whether the social evolution that Miller writes of affects aspects of languages besides their vocabularies. An argument that it does is presented in Hurford (1987), especially Ch.6.

It is clear that much of language structure can be explained by innate characteristics of the LAD; I do not claim that *all*, or even 'central' (according to some preconceived criterion of centrality) aspects of languages can be explained by factors in the Arena of Use. Bates et al. (1988:235-6) conclude: "we have found consistent evidence for 'intraorganismic' correlations, i.e. nonlinguistic factors *in the child* that seem to vary consistently with aspects of language development". Such factors belong to the Arena of Use, as defined here, but so far as is yet known, affect only development, and not the end product, the content of adult grammars. On 'extraorganismic' correlations, Bates et al. conclude: "This search for social correlates of language has been largely disappointing". (1988:236). At a global level, one should not be 'disappointed' or otherwise at how scientific results turn out. The question of interest is: '*What aspects* of language structure are attributable to the innate LAD, and what aspects to the Arena of Use?' It seems likely that the search for influences of the Arena of Use on acquired grammars will be least 'disappointing' in the marked periphery of grammar, as opposed to the core, as the core/periphery distinction is drawn by UG theorists.

2.2. Frequency, statistics and language acquisition

There is serious disagreement on the role to be played by statistical considerations in the theory of language acquisition. The tradition of learnability studies from Gold (1967), through Wexler and Culicover (1980), to such discussions as Lightfoot (1989a), assumes, but of course does not demonstrate, that statistical frequencies are totally alien to language acquisition. Theorems are derived, within a formal system, from axioms,

whose truth may perhaps be taken for granted by the inventor of the system, but which the system itself can in no way guarantee to be true. The theorems of learnability theory are derived in systems which assume a particular type of definition of 'language', in particular, languages are assumed not to have stochastic properties. But, under a different definition of 'language', different theorems are provable, showing that frequencies in the input data can be relevant to language acquisition. See, for example, Horning (1969), and comments by Macken (1987:391).

But, even with a nonstochastic definition of the adult competence acquired, it is still easily conceivable that frequency factors in the input should influence the process of acquisition. Pinker (1987), for example, assumes that adult competence is nonprobabilistic, but proposes a model of acquisition in which exposure to a piece of input data results in the 'strengths' of various elements of the grammar being adjusted, usually being incremented. The point is that in Pinker's proposal one single example of a particular structure in the input data does not automatically create a corresponding all-or-nothing representation in the child's internal grammar; it can take a number of exposures for the score on a given element to accumulate to a total of 1. Presumably, if that number of exposures isn't forthcoming in the input data, that element (rule, feature, whatever) doesn't get into the adult grammar.

Learnability theory typically operates with an assumption that the learning device is 'one-memory limited'. This is the assumption that

> "the child has no memory for the input other than the current sentence-plus-inferred-meaning and whatever information about past inputs is encoded into the grammar at that point". (Pinker, 1984:31)

But the success of learnability theory does not depend on the assumption that its 'one-memory inputs' correspond to single events in the experience of a child. It is quite plausible that there is some pre-processing front end to the device modelled by learnability theory, such that an accumulation of experiences is required for the activation of each one-memory input. Likewise it is easy to envisage that the setting of parameters in the GB/UG account needs some threshold number (more than one) of exemplars. If there were some theorem purporting to demonstrate that this is alien to language acquisition, one would need to examine carefully the relevant axioms and definitions of terms, to see if they made assumptions corresponding appropriately to data uncovered by real acquisition studies.

There *are* studies revealing relationships between acquired (albeit interim) grammars and statistical properties of the input.

"One consistent and surprising characteristic of early phonological grammars is their close relationship to frequency and distributional characteristics of not the whole language being learned but the specific input. ... (see, for example, Ingram, 1979 on French; Itkonen, 1977 on Finnish; Macken, 1980 on English and Spanish)". (Macken, 1987:385)

"... certain acquisition data in conjunction with an interpretation of the relevant evidence and correlations show that there are stochastic aspects to language acquisition, like sensitivity to frequency information". (Macken, 1987:393)

"... Gleitman et al. (1984) cite several studies showing that the development of verbal auxiliaries is affected by the statistical distribution of auxiliaries in maternal speech. In particular, mothers who produce a large number of sentence-initial auxiliaries ... tend to have children who make greater progress in the use of sentence-internal auxiliaries ... Because this auxiliary system is a peculiar property of English, it cannot belong to the stock of innate linguistic hypotheses. It follows that auxiliaries have to be picked up by some kind of frequency-sensitive general learning mechanism". (Bates et al., 1988:62)

There are several studies indicating the influence of word-frequency on internalised phonological forms.

"Neu (1980) found that adults delete the /d/ in 90 percent of their productions of *and*, compared to a 32.4 per cent rate of /d/ deletion in other monomorphemic clusters; ... Fidelholtz (1975) has observed less in the way of perceptible vowel reduction for frequent words, and Koopmans-van Beinum and Harder (1982/3) have confirmed this in the laboratory. The frequency-reducibility effect evidently holds even where syllabic and phonemic length are equated (Coker, Umeda and Browman 1973; Wright 1979), and as the effect has little to do with differences in the information content or predictability of high and low frequency words (Thiemann 1982), their different reducibility suggests that frequent (i.e. familiar) words may be stored in reduced form. [Footnote:-] Though it is not my purpose here to deal with the child's role in phonological change, my discussion here ... has an obvious bearing on this subject". (Locke, 1986:248; footnote, 524)

In the framework advanced here, either the rule deleting /d/ is not a rule of phonological competence, but belongs to the Arena of Use, or, if it is a rule of phonological competence, it is an optional rule, with applicability sensitive to factors in the Arena of Use (e.g. speed of speech)). Further evidence of a relationship between word frequency and internalised grammars is provided by Moder (1986):

"High frequency forms were found to be poorer primes of productive patterns than medium frequency forms. Furthermore, the real verb classes which showed some productivity were those with fewer high frequency forms. Because high frequency forms are often rote-learned [Bybee and Brewer, 1980], they are less likely to be analysed and related morphologically to the other members of their paradigm." (Moder, 1986:180)

Phillips (1984) discusses two distinct kinds of historical lexical phonological change, both clearly correlated, in different ways, with word-frequency:

> "Changes affecting the most frequent words first are motivated by physiological factors, acting on surface phonetic forms; changes affecting the least frequent words first are motivated by other, non-physiological factors, acting on underlying forms". (Phillips, 1984:320)

In a generative view of sound change, just as in the view I am advancing here, a sound change cannot 'act on surface phonetic forms', since what differs significantly from one generation to the next is speakers' grammars, and these contain underlying phonological forms and phonological rules, but no direct representation of surface phonetic forms. Phillips does not discuss the micro-implementation of these sound changes at the level of the individual's acquisition of language, but a straightforward interpretation of her results is as follows. Physiological factors (in the Arena of Use) produce phonetically modified forms, whose frequency gives rise, in the language-acquiring generation, to internalised underlying forms closer to the observed phonetic forms. On the other hand, non-physiologically motivated changes arise from what Phillips, following Hooper (1976), calls 'conceptually motivated change', i.e. some kind of reorganisation of the grammar for purposes of maximisation or achievement of some internal property. But these changes, apparently, cannot fly in the face of strong evidence on pronunciation coming from the Arena of Use. Only where such evidence from the Arena is very slight, as with low-frequency words, can the internal grammar reorganisation, for these cases, override the input evidence. Thus, frequency factors from the Arena of Use affect the shape of evolving languages, both positively (pressing for change) and negatively (resisting change).

The argument against the relevance of statistical considerations has another strand, which contrasts the subtlety, speed and effortlessness of our grammatical judgments with the poverty of our statistical intuitions, even the most elementary ones (this argument might cite research by Amos Tversky and Daniel Kahneman). There are several points here. Firstly, it is possible to exaggerate the subtlety, speed, and effortlessness of our grammatical judgments. Chomsky points out in many works how our grammatical knowledge needs to be 'teased out' (in the phrase used in Chomsky, 1965). For instance, "Often it is not immediately obvious what our knowledge of language entails in particular cases" (Chomsky, 1986:9), and "... it takes some thought or preparation to see that (13) has the interpretation it does have, and thus to determine the consequences of our knowledge in this case" (ibid:11).

A second point is that the relevant human frequency monitoring abilities are not poor, but quite the contrary, as a seminal publication in the psychological literature shows.

> "People of all ages and abilities are extremely sensitive to frequency of occurrence information. ... [In] the domain of cognitive psychology ... we note that the major conclusion of this area of research stands on a firm empirical base: The encoding of frequency information is uninfluenced by most task and individual difference variables. As a result, memory for frequency shows a level of invariance that is highly unusual in memory research. This is probably not so because memory is unique but because memory researchers have paid little attention to implicit, or automatic, information acquisition processes. Here we demonstrated the existence of one such process. We also showed its implications for the acquisition and utilisation of some important aspects of knowledge". (Hasher and Zacks, 1984:1385)

Hasher and Zacks also briefly discuss the relation of their work to that of Tversky and Kahneman; they conclude "... the conflict between our view and that of Tversky and Kahneman is more apparent than real" (p.1383)

Thus far, my arguments have been that statistical patterns in the input can and do affect the content of the acquired competence, perhaps especially where the language changes from one generation to the next (i.e. where the acquired competence differs from the competence(s) underlying the PLD). There is another, powerful, argument indicating the *necessity*, for language acquisition to take place at all, of a certain kind of statistical patterning in the input data. This involves what has been called the 'Semantic Bootstrapping Hypothesis', discussed in detail by Pinker (1984), but advanced in various forms by several others.

Briefly, the Semantic Bootstrapping Hypothesis states that the child makes use of certain rough correspondences between linguistic categories (e.g. Noun, Verb) and nonlinguistic categories (e.g. discrete physical object, action) in order to arrive at initial hypotheses about the structure of strings he hears. Without assuming such correspondences, Pinker argues, the set of possible hypotheses would be unmanageably large. This seems right. It is common knowledge, of course, that there is no one-to-one correspondence between conceptual categories and linguistic categories – any such correspondence is statistical. Pinker (1984:41) lists 24 grammatical elements that he assumes correspond to nonlinguistic elements. (In Pinker, 1989 the background to the hypothesis is modified somewhat, but not in any way that endangers the main point.) Now, according to the Semantic Bootstrapping Hypothesis, if these correspondences are not present in the experience of the child, grammar acquisition cannot take place.

UG theory characterises a class of possible grammars. These grammars, as specified by UG, make no mention of nonlinguistic categories. Of course, for the grammars to be usable, nonlinguistic categories must be associable with elements of a grammar. For instance, the lexical entry for *table* must, if a speaker is to use the word appropriately, get associated with the nonlinguistic, experiential concept of a table (or tablehood, or whatever).

But UG theory makes no claim about how the nonlinguistic categories are related to elements of grammars. A possible grammar, in the UG sense, might be considerably complex, and yet not contain any elements that happened to be associated with concrete physical objects, or actions, for example. And the sentences generated by such a grammar could in fact still be usable, say for abstract discourse, if, miraculously, a speaker had managed to learn it. Such a speaker could, for instance, produce and interpret such sentences as *Linguistic entities correspond roughly to non-linguistic entities*, or *Revolutionary new ideas are boring*.

But he could not talk about physical objects or actions. And if the Semantic Bootstrapping Hypothesis is true, his speech could not constitute viable input data for the next generation of learners. Thus a theory which aims to account for the perpetuation of (universals of) language across generations, via the innate LAD, actually *requires* specific conditions to be met in the Arena of Use. These conditions are not, as it happens, absolute, but are statistical.

Of course, I do not claim that statistical properties of input are the only ones relevant to the acquisition of competence. I agree with Lightfoot's position:

> "It has long been known that not everything a child hears has a noticeable or long-term effect on the emergent mature capacity; some sifting is involved. Some of the sifting must surely be statistical, some is effected through the nature of the endowed properties ..." (Lightfoot, 1989b:364)

Facts of grammar are likely to be distributed along a dimension according to whether their acquisition is sensitive to frequency effects in the input data. Some aspects of grammar may involve very rapid fixing (once the child is 'ready') on the basis of very little triggering experience. Other aspects of grammar may be harder to fix, requiring heavier pressure (in the form of frequency, among other things) from the input experience. This suggested dimension is a graded version of Chomsky's binary core/periphery distinction.

Chomsky seems to acknowledge the greater role of input data for the acquisition of the periphery of grammar:

> "... we would expect phenomena that belong to the periphery to be supported by specific evidence of sufficient 'density'..." (Chomsky, 1986:147)

Whether or not Chomsky intended frequency considerations to contribute to this 'density', there is no principled reason why they should not. As Pinker writes:

> "Ultimately no comprehensive and predictive account of language development and

> language acquisition can avoid making quantitative commitments altogether. After all, it may turn out to be *true* that one rule is learned more reliably than another only because of the steepness of the relevant rule strengthening function or the perceptual salience of its input triggers". (Pinker, 1984:357)

Pinker then states a methodological judgement that 'For now there is little choice but to appeal to quantitative parameters sparingly'. I share his apprehension about the possibility of 'injudicious appeals to quantitative parameters in the absence of relevant data', but the solution lies in making the effort to obtain the relevant data, rather than in prejudging the nature (statistical or not) of the theories that are likely to be correct.

2.3. Grammaticalisation, syntacticisation, phonologisation

Previous work has identified a phenomenon of 'grammaticalisation', dealing precisely with historical interactions between the Arena of Use and individual linguistic competences. Some such work is vitiated by a misguided attempt to abolish the competence/performance distinction.

Givon (1979:26-31) surveys a number of cases in which, on one view of grammar (a view Givon appears emphatically not to hold) "... one may view a grammatical phenomenon as belonging to the realm of competence in one language and performance-text frequency in another" (26). Givon's examples are: (i) the definiteness of subjects of declarative clauses, obligatory in some languages, but merely preferred in others; (ii) the definiteness of referential objects of negative sentences, obligatory in some languages, but merely preferred in others; and (iii) the lack of an overt agent phrase with passive constructions, obligatory in some languages, but merely the preferred pattern in others. The preferences involved can be very strong, but in the languages where the facts seem not to be a matter of absolute rule, but of preference, one can find isolated examples of the pattern that would be ungrammatical in the other language.

In precisely similar vein, though not sharing Givon's conclusions, Hyman (1984) writes that he has been

> "... intrigued by a puzzling recurrent pattern which can be summarized as in
>
> (1) a. Language A has a [phonological, phrase-structure, transformational] rule R which produces a discrete (often obligatory) property P;
> b. Language B, on the other hand, does not have rule R, but has property P in some (often nondiscrete, often nonobligatory) less structured sense". (Hyman 1984:68)

And Corbett (1983) in an impressively documented study gives many instances where one Slavic language has an absolute rule which is paralleled

by a statistical tendency in some other Slavic language. One such case is:

> The agreement hierarchy
>
> attributive - predicate - relative pronoun - personal pronoun
>
> "In absolute terms, if semantic agreement is possible in a given position in the hierarchy, it will also be possible in all positions to the right. In relative terms, if alternative agreement forms are available in two positions, the likelihood of semantic agreement will be as great or greater in the position to the right than in that to the left." (Corbett, 1983:10-11)

Givon offers an alternative view to the one quoted above:

> "Or one may view the phenomenon in both languages in the context of 'communicative function', as being essentially *of the same kind.* The obvious inference to be drawn from the presentation is as follows: If indeed the phenomenon is of the same kind in both languages, then the distinction between competence and performance - or grammar and speaker's behaviour - is (at least for these particular cases) untenable, counterproductive, and nonexplanatory." (Givon, 1979:26)

This passage, like other polemical passages in linguistics, is a curious mixture of over- and understatement. It ends like a Beethoven symphony, with repeated heavy chords, slightly varied, but united in their effect 'untenable, counterproductive, and nonexplanatory'. But immediately before is the weakening parenthetical caveat '(at least for these particular cases)', and the whole conclusion is in fact embedded in a conditional, '*If indeed* the phenomenon *is* of the same kind in both languages' [emphasis added, JRH]. So, if the condition is not met and the phenomena are not of the same kind in both languages, the three big guns 'untenable, counterproductive, and nonexplanatory' aimed at the competence/performance distinction don't actually go off. And, even if the condition is met, they may only be aimed at the distinction 'for these particular cases'. Much of Givon's book reflects this kind of rhetorical mixture. The message, if interpreted as urging alternative *emphases* in linguistic study, is entirely reasonable; more work *should* be done on the relation between communicative, pragmatic, and discourse phenomena and grammar, and this, thanks to the efforts of people like Givon, is beginning to happen. Clark and Haviland, 1974 is another work in which a reasonable argument for emphasis on discourse study is in places rhetorically inflated to a claim that "the borderline between the purely linguistic and the psychological aspects of language ... may not exist at all". (p.91).) But the relation between grammar on the one hand and discourse phenomena on the other cannot be studied if the two sets of phenomena actually turn out to be the same

thing, as Givon appears in places to believe. There are good reasons to maintain a distinction between facts of grammar and facts of discourse, and Givon manages throughout his book to write convincingly as if the distinction were valid. What is of great interest is the parallelism between the two domains, illustrated by Givon, Hyman, and Corbett, as cited at the beginning of this section.

Absolute grammatical rules in one language paralleled closely by statistical discourse preferences in another language may seem something of a puzzle. But the puzzle can be relatively easily resolved. Let me risk giving a nonlinguistic analogy, asking the reader to make the usual mutatis mutandis allowances necessary for all analogies.

Some people eat a variety of foods, but, without having made any decision on the matter, happen hardly ever to eat meat; other people are vegetarians by decision, though sometimes they may accidentally eat meat. Some people, as a matter of habit, drink no alcohol; for others, this is not a matter of habit, but of principle. Some people are pacific by nature; others are pacifists on principle. The principled vegetarians, teetotallers and pacifists have made conscious absolute decisions which are parallel to the statistical behavioural tendencies of certain other people. But there is a valid distinction to be made between the two categories. This distinction is not particularly obvious from mere observation of behaviour. But as humans, we have the benefit of (some) self-knowledge, and we *know* that there is a difference between a principled vegetarian and a person who happens hardly ever to eat meat, and between a principled pacifist and a pacific person.

No analogy is entirely apt, however. This one suffers in at least two ways. Firstly, speakers of a language do not normally make conscious decisions, like the pacifist or the vegetarian, about their own rules of grammar; and secondly, the vegetarian/teetotaller/pacifist analogy distinguishes between individuals in the same community, whereas rules of grammar tend to be shared by members of a speech-community. What I hope this analogy demonstrates is that similar overt patterns of behaviour can be attributed to different categories, such as fact of discourse, or fact of grammar. The categories may also be historically related, as I believe discourse and grammar are, but they are not now a single unified phenomenon.

I assume, then, that there is factual content to the notion of following an internalised rule. Chomsky (1986), in a lengthy and cogent discussion, disposes quite satisfactorily, in my view, of the Wittgensteinian objections, taken up by Kripke (1982), to attribution of rule-following by other organisms, be they fellow-speakers of one's language, foreign humans, or even other animals. Wittgensteinians (among whom one would include, for instance, Itkonen, 1978) have often objected to the generativists'

interpretation of rules of language as essentially belonging to the psychology of individuals, and a generativist response to this theme of Wittgenstein's is now satisfactorily articulated. (Perhaps the delay in responding arose from the enigmatic style of Wittgenstein's original presentation, and Kripke's reformulation gave the clarity needed for a careful rebuttal.) I uphold the view that rules of language belong to individual psychology, and should not in any collective sense be attributed to communities (e.g. as social norms). Thus far, I agree completely with Chomsky's position on rules and rule-following. But now here is where we part company: "reference to a community seems to add nothing substantive to the discussion" (Chomsky, 1986:242). I maintain, on the contrary, that communities play a role in determining what rules an individual acquires (which is obvious), and, more generally, that general facts about human communal life play a role in determining the *kinds* of rules that individuals born into any human community acquire. Pateman expresses the idea so well that his words are worth repeating:

> "... through time the content of mentally represented grammars, which are not in my view social objects, comes to contain a content which was in origin clearly social or cultural in character." (Pateman, 1985:51)

The historical mechanism by which facts of discourse 'become' facts of grammar is often labelled 'grammaticalisation'. To prevent confusion, it should be stressed that the result of this process does not necessarily involve a class of previously ungrammatical strings becoming grammatical. The converse process can also occur. What gets grammaticalised is a *pattern*, or configuration of facts, not some class of strings which happens to participate in such a pattern. The following are the main interesting possibilities, in terms of two classes of strings, A and B, which are in some sense functionally equivalent (e.g. (partially) synonymous).

(5) A and B are both grammatical, but A is preferred in use.

Diachronic change in either direction ↕

A is grammatical, and B is ungrammatical, though B may occur in use

Change in either direction involves a new fact of grammar emerging, which is why such changes are aptly called 'grammaticalisation'. But only change in one direction (upward in (5)) involves previously ungrammatical strings becoming grammatical. Change in either direction would account for the

parallelisms noted by Givon, Hyman, and Corbett. Another possibility is:

(6) A and B, both grammatical, are wholly equivalent in meaning and use.

Diachronic change in either direction. ↕

A and B, both grammatical, but differ slightly in meaning and use.

As the relation of surface forms to their linguistic meanings is a matter of grammatical competence, this is also a case of the emergence of a new fact of grammar, and aptly called 'grammaticalisation'.

How does the mechanism of grammaticalisation work and how does it relate to the question of nativist versus functional explanations? I beg leave to quote myself.[4]

> "In the model proposed, individual language learners respond in a discrete all-or-nothing way to overwhelming frequency facts. Language learners do not merely adapt their own usage to mimic the frequencies of the data they experience. Rather, they 'make a decision' to use only certain types of expression once the frequency of those types of expression goes beyond some threshold. At a certain point there is a last straw which breaks the camel's back and language learners 'click' discretely to a decision about what for them constitutes a fact of grammar. What I have in mind is similar to Bally and Sechehaye's suggestion about Saussure's view of language change. 'It is only when an innovation becomes engraved in the memory through frequent repetition and enters the system that it effects a shift in the equilibrium of values and that language [langue] changes, spontaneously and *ipso facto*' (Saussure, 1966:143n). Bever and Langendoen (1971:433) make the same point nicely by quoting *Hamlet*: 'For use can almost change the face of nature' ". (Hurford, 1987:282-3, slightly adapted)

Beyond the kind of vague remarks cited above, no-one has much idea of how grammaticalisation works. Givon's book documents a large number of interesting cases, but his account serves mainly to reinforce the conclusion *that* grammaticalisation happens, rather than telling us *how* it happens. And of course the fact that it does happen, that aspects of performance get transmuted into aspects of competence, reinforces, rather than undermines, the competence/performance distinction. But one thing that is clear about grammaticalisation is that the LAD plays a vital part. This emerges from Givon's discussion of Pidgins and Creoles, in which the discrete step from Pidgin to Creole coincides with language acquisition by the first-generation offspring of Pidgin speakers.

> "Briefly, it seems that Pidgin languages (or at least the most prevalent type of Plantation Pidgins) exhibit an enormous amount of internal variation and inconsistency both within the output of the same speaker and across the speech community. The variation is massive to the point where one is indeed justified in asserting that the Pidgin has *no stable syntax*. No consistent "grammatical" word-order can be shown in a Pidgin, and little or no use of grammatical morphology. The rate of delivery is excruciatingly slow and halting, with many pauses. Verbal clauses are small, normally exhibiting a one-to-one ratio of nouns to verbs. While the subject-predicate structure is virtually undeterminable, the topic-comment structure is transparent. Virtually no syntactic subordination can be found, and verbal clauses are loosely concatenated, usually separated by considerable pauses. In other words, the Pidgin speech exhibits almost an extreme case of the *pragmatic* mode of communication.
>
> In contrast, the Creole – apparently a synthesis *di novo* [sic] by the first generation of native speakers who received the Pidgin as their data input and proceeded to "create the grammar" – is very much like normal languages, in that it possesses a syntactic mode with all the trimmings ... The amount of variation in the Creole speech is much smaller than in the Pidgin, indistinguishable from the normal level found in "normal" language communities. While Creoles exhibit certain uniform and highly universal characteristics which distinguish them, in degree though not in kind, from other normal languages, they certainly possess the entire range of grammatical signals used in the syntax of natural languages, such as fixed word order, grammatical morphology, intonation, embedding, and various constraints". (Givon, 1979:224)

This passage makes the case so eloquently for the existence of an innate Language Acquisition Device playing a large part in determining the shape of normal languages that one would not be surprised to find it verbatim in the introduction to a text on orthodox Chomskyan generative grammar. In my terms, the prototypical Pidgin is a hybrid monstrosity inhabiting the Arena of Use, limping along on the basis of no particular shared core of individual competences. The main unifying features it possesses arise from its particular spatial/temporal/social range in the Arena of Use. When a new generation is born into this range, and finds this mess, each newborn brings his innate linguistic faculty to bear on it and helps create, in interaction with other members of the community, the grammar of the new Creole.

The picture just given is, by and large, that of Bickerton's Language Bioprogram Hypothesis (Bickerton, 1981), and is probably correct in broad outline, if no doubt an oversimplification of the actual facts. "Usually, however, the trigger experience of original creole speakers is shrouded in the mists of history, and written records of early stages of creole languages are meagre." (Lightfoot, 1988:100) A vast amount of empirical research into the creolisation process needs to be done before interesting details become discernible, but clearly the focal point of the process is the point where the innate LAD meets the products of the Arena of Use. The step from a Pidgin to a Creole is an extreme case of many simultaneous

grammaticalisations across virtually the whole sweep of the (new) language. Creolisation is massive grammaticalisation. But it is also, due to the historical rootlessness of the Pidgin, grammaticalisation with a very free hand. The LAD can impose its default values against weak opposition from the Pidgin PLD. In discussing grammaticalisation, I do not presuppose that the input to the process is *necessarily* some pattern evident in use. My position is that grammaticalisation is the creation, by the LAD, of new facts of grammar. Where the input is chaotic, the LAD has a very free hand, and the new facts of grammar reflect the LAD's influence almost solely. But where patterns of use exist in the input data, the new facts of grammar may in certain instances reflect those patterns. We can call these latter cases 'grammaticalisations of patterns of use', and the former (dramatic creole) cases 'grammaticalisations from nothing'.

Creoles are in some sense more natural than languages with long histories. Languages with long histories become encrusted with features that require non-default setting from the LAD, and even rote-learning. These encrustations are due to innovations in the Arena of Use over many generations. Many of these developments can be said to be functionally motivated. I have already mentioned in passing several historical studies (Bever and Langendoen, 1971, Phillips, 1984) which make at least prima facie cases for the influence, across time, of use on structure. And in section 2.6, I will add to the list of recent historical linguistic studies which point to the role of the Arena of Use in determining, at least in part, the contents of grammatical competence. In these cases, the languages have drifted, due to pressures of use, to become, in some sense, historically more 'mature' than a new creole.

It seems reasonable to suppose that sheer statistical frequency of particular patterns in the Arena of Use will play some part in determining what grammatical rules will be formed. This is one way in which a parallelism between discourse patterns and grammatical rules would arise. But of course the LAD is not merely quantitatively, but also qualitatively selective. It is not the case that any, i.e. every, frequent pattern becomes grammaticalised. If this were so, the most common performance errors, hesitation markers and such like would always get grammaticalized, which of course they often don't. (But note that hesitation markers do tend to become fitted into the vowel system of the dialect in question, i.e. to become phonologised. Cf. the various hesitation vowels in RP ([3:]), Scots English ([e:]), and French ([ø:]).) I believe that Lightfoot, in his 1988 paper, somewhat oversimplifies the relation between the qualitative and the quantitative selectivity of the LAD in the following remarks:

> "The most obvious point is that not everything that the child hears 'triggers' a device in the emerging grammar. For example, so-called 'performance errors' and slips of the tongue do not entail that the hearer's grammar be amended in such a way as to generate such deviant expressions, presumably because a particular slip of the tongue does not occur frequently enough to have this effect. This suggests that a trigger is something that is robust in a child's experience, occurring frequently. Children are typically exposed to a diverse and heterogeneous linguistic experience, consisting of different styles of speech and dialects, but only those forms which occur frequently for a given child will act as triggers, thus perpetuating themselves and being absorbed into the productive system which is emerging in the child, the grammar." (Lightfoot, 1988:98)

This seems to equate 'potential trigger experience' with 'frequent experience'. Lightfoot has now developed his ideas on the child's trigger experience further (Lightfoot, 1989), but he still holds that some statistical considerations are relevant. While, with Lightfoot, I believe that frequency in the Arena of Use is an important determinant of the grammars that children acquire, there must also be substantial qualitative selectivity in the LAD. Some aspects of competence can be picked up on the basis of very few exemplars, while the LAD stubbornly resists acquiring other aspects for which the positive examples are very frequent. The particular qualitative selectivity of the LAD is what is studied under the heading of grammatical universals, or UG.

2.4. The role of invention and individual creativity

Prototypical short-term functional explanations involve the usefulness of some aspect of a language making itself felt within the time a single individual takes to acquire his linguistic competence (although I shall later mention a version of the same basic mechanism that happens to take somewhat longer). This period may vary from a dozen years, for grammatical constructions, to a whole lifetime, for vocabulary. But, in the prototypical case, a short-term functional explanation involves postulating that each individual acquiring some language recognizes (perhaps unconsciously) the usefulness of some linguistic element (word, construction, etc.) and adds that item to his competence *because it is useful.* Some universal facts of vocabulary, such as the fact that every human language has at least one word with a designatum in the water/ice/sea/river area, can be illuminated in this way, as can also many language-particular facts, such as those of color, plant, and animal taxonomies worked on in detail by the 'ethnographic semantics' movement (e.g. Brown, 1984). Thus, those aspects of languages for which short-term functional explanations are available are characteristically transmitted culturally. Individuals actually learn these aspects of their language from other members of their com-

munity. They are not innate. Such aspects of languages, therefore, are typically well-determined by the observable data of performance, since they need to be sufficiently obvious to new generations to be noticed and adopted.

Obviously, quite a lot is innate in the lexicon too. For instance, no single verb can mean 'eat plenty of bread and...', 'persuade a woman that...', 'read many books but not...'. The constraints on possible lexical meanings are strong and elaborate. My point is that, *within* such innately determined constraints, the matter of what lexical items a language possesses is influenced by factors of usefulness. Individual inventiveness cannot violate the innately determined boundaries, see Hurford, 1987, Ch.2,Sec.5, for a detailed discussion of the relation of individual inventiveness to the capacity for language acquisition.

Aspects of languages transmitted culturally from one generation to the next because of their usefulness have their origins in the inventiveness and creativity (presumably in some sense innate) of the individuals who first coined them and gave them currency. In the field of vocabulary again, it is uncontroversial that new words are invented by individuals, or arise somehow from small groups. Often it is not possible to trace who the first user of a new word was, but nevertheless there must have been a first user. In other parts of languages, such as their phonological, morphological, syntactic, semantic and pragmatic rule components, it is difficult to attribute the origins of particular rules to the creativity of individuals or groups, but even here a kind of attenuated creativity in the use of language, proceeding by small increments over many generations, seems plausible. The approximate story would be of existing rules having their domain of application gradually extended or diminished due to a myriad of small individual choices motivated by considerations of usefulness. Very few rules of syntax are completely general in the sense of having no lexical exceptions. Such sets of lexical exceptions are augmented or lessened continually throughout the history of languages. The specifically functional considerations, that is considerations of usefulness, which motivate such changes in the grammar of a language are of course usually impossible to identify with accuracy, and will remain so until we have much subtler theories and taxonomies of language use (which will help us to define the notion of usefulness itself more precisely).

The historical role of invention and creativity that I have in mind is envisaged by Gropen et al. (1989) and described by Mithun (1984):

> "Instead, it could be that the historical processes which cause lexical rules to be defined over some subclasses but not others seem to favour the addition or retention of narrow classes of verbs whose meanings exemplify or echo the semantic structure created by the rule most clearly. The full motivation for the dativisability of a narrow class may

> come from the psychology of the first speakers creative enough or liberal enough to extend the dative to an item in a new class, since such speakers are unlikely to make such extensions at random. Thereafter speakers may add that narrow class to the list of dativisable classes with varying degrees of attention to the motivation provided by the broad-range rule – by recording that possibility as a brute memorised fact, by grasping its motivation with the aid of a stroke of insight recapitulating that of the original coiners, or by depending on some intermediate degree of appreciation of the rationale to learn its components efficiently, depending on the speaker and the narrow class involved". (*Gropen et al.*, 1989:245)

> "But in Mohawk, where NI [= noun incorporation] of all types is highly productive, speakers frequently report their pleasure at visiting someone from another Mohawk community and hearing new NI's for the first time. They have no trouble understanding the new words, but they recognise that they are not part of their own (vast) lexicon. When they themselves form new combinations, they are conscious of creating 'new words', and much discussion often surrounds such events." (Mithun, 1984:889)

The acts of individual speakers in responding creatively to considerations of usefulness are analogous to micro-events at the level of molecules, and the large movements of languages discernible to historical linguists are analogous to macro-events, such as those described in geophysical terms of plate tectonics (this analogy is Bob Ladd's). Whether or not we call a language in which there has been one micro-change a different language is a question of terminology. Let us adopt, temporarily and for argument's sake, the rigid convention that any one change, however slight, in a language L_n produces a different language L_{n+1}. This effectively equates 'language' with some abstraction even lower in level than 'idiolect', and so is not a generally useful convention in talking about language[5]. But, adopting this usage, competition in the Arena of Use determines whether L_n or L_{n+1} survives. These minimally differing languages may continue to co-exist, because neither is significantly more useful than the other, or one may replace the other because it is in some sense more useful. Adopting a different terminological convention, wherein 'languages' are grosser entities, distinguished by masses of detailed differences, it is still competition in the Arena of Use which decides the survival of languages. The 'languages' I have in mind in this paragraph are I-languages. But since they, existing only inside speakers, can never come into contact with each other, the competition between them is actually fought out through the medium of their corresponding E-languages in the Arena of Use. (An approximate analogy would be a tournament acted out by marionette puppets whose behavioural repertoires (kick, punch, etc.) are specified by different programs of their robot operators, though the set of programs available in principle to all robots is the same. When a puppet loses a match, the program in the robot that was running it is eliminated. But remember that no analogy is perfect.)

A schematic representation of the state of affairs postulated in a functional explanation of the short-term type is given in Figure 5, below. Note that the 'languages' mentioned in this diagram are E-languages, since they exist in and through the Arena of Use; that is, they correspond to the competing marionettes of the analogy of the previous paragraph.

SHORT-TERM MECHANISM OF FUNCTIONALLY MOTIVATED CHANGE ONTOGENETIC (OR GLOSSOGENETIC) MECHANISM

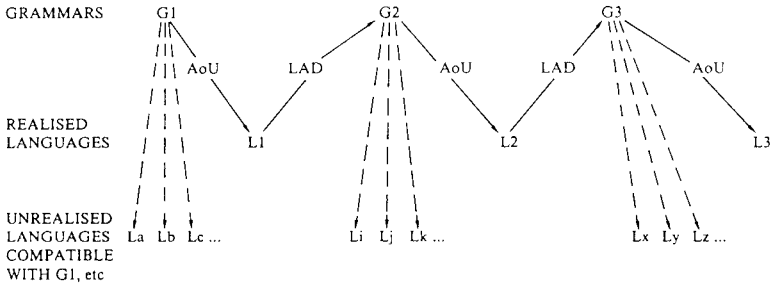

Fig. 5.

The upper two levels in this diagram indicate the course of actual linguistic history: the actually mentally represented grammars G1, G2, G3, ..., and the actually realised languages L1, L2, L3, ... The bottom level in the diagram represents alternative language histories – what languages might have been realised if the pressures of the Arena of Use had been other than what they actually were. These possible but unrealised languages can be thought of as aborted due to competition in the Arena of Use from a more successful rival language. Competition in the Arena of Use, in the case of this short-term functional mechanism, is therefore between possible languages defined by the *same* LAD. (Figure 5 is in fact another variant of Andersen's scheme in Figure 4.) The unrealised languages are possible but non-occurring aggregates of real speech events in the language community, alternative courses of history, in effect.

The scheme shown in Figure 5 is obviously idealised in many ways. One aspect of this idealisation worth mentioning is the fact that only one LAD is represented at any transition, whereas in fact language change is mediated by whole populations of LADs (tokens not types), all (1) exposed to different (though partially intersecting) data, (2) possibly themselves subject to some maturational change (see White, 1982:68-70, Borer and Wexler, 1987, 1988), and (3) perhaps even not originally completely uniform. In a real case, some individuals would internalise grammars slightly different from those

internalised by others. This difference would be reflected by statistical changes in the Arena of Use, which in turn might prompt a rather larger proportion of language learners in the next generation to acquire grammars with a certain property. In this way, it might take many generations for a whole population to accomplish what with historical hindsight looks like a single discrete change. The term 'ontogenetic mechanism' might well be reserved for a case where a whole nonstatistical language change is achieved in a single generation, rather like the Bickerton/Givon picture of the leap from Pidgins to Creoles. That is, the new (version of the) language grows, fast, in just the time it takes one generation of individuals to acquire/create it. The slower version of the mechanism, which takes more than one generation, could appropriately be called the 'glossogenetic mechanism'. The only difference between the ontogenetic and the glossogenetic mechanism is in the number of generations taken.

2.5. The problem of identifying major functional forces

This picture of functionally motivated language change has its opponents. One of the fiercest and most sustained critiques of this general point of view that I am aware of is in Lass (1980:64-97). Lass's view (in which he is not alone) is summed up in:

> "Merely on the evidence provided so far, if my arguments are sound, the proponents of any functional motivation whatever for linguistic change have to do one of two things:
>
> (i) Admit that the concept of function is *ad hoc* and particularistic and give up; or
> (ii) Develop a reasonably rigorous, non-particularistic theory with at least some predictive power; not a theory based merely on *post hoc* identification plus a modicum of strategies for weaseling out of attempted disconfirmations.
>
> This is the picture as I see it: (i) is of course the easy way out, and (ii) seems to be the minimum required if (i) is not acceptable. I am myself not entirely happy with (i), and it should probably not be taken up – though failing a satisfactory response to (ii) it seems inevitable." (Lass, 1980:79-80)

Lass discusses functional explanation under three subheadings: 'preservation of contrast', 'minimization of allomorphy', and 'avoidance of homophony', and convincingly demolishes claims by various scholars to have explained particular historical linguistic changes in such 'functional' terms. But in fact these attempted explanations are not genuinely functional according to the spirit in which I have argued the term should be taken. It is crucial to note that 'contrast', 'allomorphy', and 'homophony', as Lass uses them, are terms describing a language system, and not language

use. In other words, quite clearly, these terms do not describe phenomena in the Arena of Use. Instances of contrast, mean degree of allomorphy, and pervasiveness of homophony can all be ascertained from inspection of a grammar, without ever observing a single speaker in action. This is of course what makes them attractive to many linguists. These are formal properties, in the same way that the simplicity of a grammar, measured in whatever way one chooses, is a formal property. Martinet's 'functional load' is likewise a formal property of language systems, not of language use, which may account for the failure of that concept to blossom as a tool of functional explanation. Obviously, the presence of contrast makes itself *felt* in the Arena of Use, but then so do most other aspects of grammars.

In fact, an old and important debate in the transition from post-Bloomfieldian structuralist phonology to generative phonology sheds light on the relation between contrast, competence, and functionally motivated language change. The classical, taxonomic, or autonomous phoneme, whose essence was that it was defined in terms of contrast, was the central concept of pregenerative phonology. This was before the emergence of a better understanding of the competence/performance, or I-language/E-language, distinction, that came with the advent of generative linguistics. To the surprise of some, it turned out that generative phonology, conceived as a model of an individual's mentally represented knowledge of the sound pattern of his language, had no place at all for the classical phoneme. The classical phoneme simply did not correspond to any linguistically significant level of representation in competence grammars. The phonemicists who found this puzzling had no arguments against this conclusion, yet puzzlement remained, in some quarters. And, in 1971, a postscript to the debate appeared, an article by Schane (Schane, 1971), which pointed the way to a resolution of the puzzle. But even 1971 was too close to the events for matters to have become completely clear, and Schane's postscript still leaves something rather unsettled; I now offer a post-postscript, taking Schane's ideas, and showing how they can be well accommodated within the picture of the interaction between the LAD and the Arena of Use.

Schane points to attested or ongoing sound changes in a number of languages (French nasalisation, Rumanian Palatalisation, Rumanian de-labialisation, Nupe palatalisation and labialisation, and Japanese palatalisation). These changes conform to a pattern:

> "If, on the surface, a feature is contrastive in some environments but not in others, that feature is lost where there is no contrast". (Schane, 1971:505)

On the basis of these examples, Schane maintains that, for the speakers involved, the (approximately) phonemic level of representation at which

these contrasts exist must have had some psychological validity. But he has this problem:

> "Transformational phonology rejects the phoneme as a unit of surface contrast, so the theory has no way of identifying contrasts, and therefore no basis for identifying *alternations* (cf. Schane 1971:514). No point in derivations exists where contrasts are identified". (Hudson, 1980:116)

Faced with the problem of reconciling some kind of psychological validity for the phoneme with the accepted conclusions of generative phonology, Schane argues in detail that representations at the phonemic level can be *calculated* from generative descriptions. The necessary calculations involve a partition of the rules into two types (morphophonemic and phonetic) and inspection of the derivations involving just rules of the former type. Note that the partition of phonological rules into morphophonemic and phonetic is also not directly represented in a generative grammar (of the type Schane was assuming) and must itself be calculated. So though a phonemic level may be accessible through a generative grammar, it is certainly not retrievable in any simple way. Schane's dilemma was that he, like others, "felt guilty about disinheriting the child [the phoneme]" (520), but since linguistic theories at the time were only competence theories, he had no obvious place to locate the phoneme.

The classical phoneme was never as well-behaved as its structuralist proponents, some of whom wanted to build it into a bottom-up discovery procedure for grammars, would have liked. Languages often use a contrast distinctively in one environment, but ride roughshod over the distinction in productive phonological rules elsewhere. An example is English /s/~/z/, a phonemic contrast 'demonstrated' by the existence of many minimal pairs (*sue/zoo, bus/buzz, racer/razor*), but neutralized in many environments by some of the most productive phonological rules of English, the voicing assimilation rules involving the plural, 3rd person singular present tense, and possessive morphemes. Naturally, the phonemicists had a story to tell about such problems, but they were typically epicyclic. What could not be saved was the idea that the main thing a speaker knows about the sounds of his language is a set of surface contrasts, which serve everywhere to 'keep words apart' (Hockett's phrase).

But of course, by and large, in the rough and tumble of everyday communication, enough words do get kept apart for decoding and successful communication to take place, much of the time. If phonological rules could obliterate all predictable distinctions between words, communication would break down. Some neutralizations are clearly permissible; the typical redundancy of language allows decoding in spite of them. But the situation cannot get out of hand. This suggests that the proper place for something

like the 'Phonemic principle' is the Arena of Use. Speakers who allow their phonetic performance to stray too far away from the surface contrasts used as clues in reception by hearers are likely to be misunderstood. To remain as (linguistically) successful members of the speech community, they learn to respect, in a rough and ready way, a degree of surface contrastivity.

I believe that Schane's basic account of the sound changes he discusses does illuminate them. Something puzzling (e.g. denasalisation following hard on the heels of nasalisation) is made to seem less puzzling by drawing attention to the fact that this happened in an environment where no surface contrast was lost. But Schane's principle is only explanatory in this weak sense; it lacks the predictive power that Lass calls for, and falls into Lass's category of 'a theory based merely on *post hoc* identification'. As Hock (1976) points out:

> "Though such changes undeniably occur, [Schane's] general claim is certainly too strong. Note, first of all that the similar loss of *u*-umlaut before remaining *u*, referred to as an 'Old Norse' change ..., is actually limited to Old Norwegian (cf. Benediktsson 1963) – Old Icelandic does not participate in it: ... Moreover, among such frequent conditioned changes as palatalization and umlaut, examples of such a 'reversal' of change seem extremely infrequent, suggesting that the phenomenon is quite rare". (Hock, 1976:208)

What is needed, to explain particular sound changes, is a demonstration that particular contrasts are felt so important that actions occur in the Arena of Use tending to prevent loss of such contrasts. Such demonstrations are likely to be very difficult, because they involve delving into the very messy data of the Arena of Use in search of clear indications involving individual words, phonemes, etc. The confrontation with the messy data of the Arena of Use is, however, far less daunting if one heeds the crucial point made by Foley and Van Valin:

> "It must be emphasised that functional theories are *not* performance theories. That is, they seek to describe language in terms of the *types* of speech activities in which language is used as well as the *types* of constructions which are used in speech activities. They do not attempt to predict the actual *tokens* of speech events. ... They are theories of systems, not of actual behavior". (Foley and Van Valin, 1984:15)

Unfortunately, this is expressed slightly inaccurately, in my terms. I would rather have said: 'functional theories are theories of performance types, and not of performance tokens'. The point is clear, however, and should be invoked to protect functional theories from disappearing without trace into the ultimate morasse of particular events. But the warning may still not be strong enough, because even functional hypotheses in terms of

particular construction types and speech activity types are likely to be met by counterexamples.

To overcome this and Lass's correct criticism of the 'particularism' of functional explanations, we somehow need to get a good *statistical* grip on the functional factors that affect language change. It is to be hoped that broad classes of events in the Arena of Use are susceptible to statistical treatment, even though individual events may appear to be more or less random. A theory of functional language change is, for the foreseeable future, only likely to be successful in characterising the statistical distribution of possible end-results of change. In this way it will be predictive in the same sense as, say, cosmology, is predictive. A given cosmological theory may predict that background microwave radiation from all directions in the universe varies only within very narrow limits (a statistical statement), but it will make no predictions at all about the particular variations.

In starting to get to theoretical grips with phenomena in the Arena of Use, it will be important to note Bever's guiding words:

> "I have attempted to avoid vague reference to properties such as "mental effort" "informativeness" "importance" "focus" "empathy" and so on. I do not mean that these terms are empty in principle: however they are empty at the moment, and consequently can have no clear explanatory force". (Bever, 1975:600-601)

Many well-intentioned attempts to establish foundations for functional theories of language change, as, for instance, in Martinet (1961), fall foul of this problem. But there are positive developments, too. The parsing explanation for word order universals offered by Hawkins (1990) makes precise a notion of economy in parsing that rescues a 'principle of least effort' (Zipf, 1949), in this area at least, from vagueness and vacuity. And I would add a reservation to Bever's warning. Terms and concepts acquire explanatory force by being invoked in plausible explanations of wide ranges of phenomena. We don't know in advance just where on the theoretical/observational continuum notions like 'mental effort' and 'informativeness' will fall. They may turn out to be relatively abstract notions, embedded in a quite highly structured theory. In such a case, their explanatory force would derive from the part they play in the explanatory success of the theory as a whole: it will not be possible to evaluate their contribution in isolation.

Lass's reluctance to take up his constructive second option, 'Develop a reasonably rigorous, non-particularistic theory with at least some predictive power', is curious. We build theories, the best that the domains concerned permit, to gain illumination about the world. As long as we don't, we remain in the dark. Of course, we should also avoid building theories only where the (usually mathematical) light is good, like the

proverbial man searching for his keys under the street lamp, rather than where he had dropped them, because the light was better under the lamp. But it is precisely because the light is (at present) dim in the area of functional influences on language change that adequate functional theories have not emerged.

Perhaps in some cases there are indeed no functional causes of language change, and the changes merely come about by random drift such as one may expect in any complex culturally transmitted system. But it would be quite unreasonable to assert that in no cases does the factor of usefulness exert a pressure for change. The fact that we are unable to pinpoint specific instances should not be confused for an argument that changes caused by factors of usefulness do not exist. We can't see black holes in space, but we have good reasons to believe they exist. Does anyone really doubt that languages are useful systems and that (some) changes in them are brought about by factors of usefulness? The only (!) issue is of the precise nature and extent of the mechanisms involved.

2.6. Language drift

A number of recent studies in diachronic linguistics have proposed evolutionary tendencies in the histories of languages.

Bybee (1986), for example, argues for the universal origin of grammatical morphemes in independent lexical items.

> "... the types of change that create grammatical morphemes are universal, and the same or similar material is worn down into grammatical material in the same manner in languages time after time ..." (Bybee, 1986:26)

> "... grammatical morphemes develop out of lexical morphemes by a gradual process of phonological erosion and fusion, and a parallel process of semantic generalisation". (Bybee, 1986:18)

Mithun (1984) proposes that noun incorporation (NI) develops diachronically along a specific route:

> "NI apparently arises as part of a general tendency in language for V's to coalesce with their non-referential objects, as in Hungarian and Turkish. The drift may result in a regular, productive word formation process, in which the NI reflects a reduction of their individual salience within predicates (Stage I). Once such compounding has become well established, its function may be extended in scope to background elements within clauses (Stage II). In certain types of languages, the scope of NI may be extended a third step, and be used as a device for backgrounding old or incidental information within discourse (Stage III). Finally, it may evolve one step further into a classificatory

> system in which generic NP's are systematically used to narrow the scope of V's with and without external NP's which identify the arguments so implied (Stage IV)". (Mithun, 1984:891)

Mithun goes on to describe other tendencies for change that languages may undergo, in cases where the evolutionary process is arrested at any of these stages.

Traugott (1989) discusses 'paths of semantic change' in terms of the following three closely related tendencies:

> "Tendency I: Meanings based in the external described situation > meanings based in the internal (evaluative/perceptual/cognitive) described situation.
>
> Tendency II: Meanings based in the external or internal described situation > meanings based in the textual and metalinguistic situation.
>
> Tendency III: Meanings tend to become increasingly based in the speaker's subjective belief state/attitude toward the proposition. ...
>
> All three tendencies share one property: the later meanings presuppose a world not only of objects and states of affairs, but of values and of linguistic relations that cannot exist without language. In other words, the later meanings are licensed by the function of language". (Traugott, 1989:34-35)

Naturally, the proposals of Bybee, Mithun, and Traugott are subject to normal academic controversy, but it seems likely that some core of their central ideas will stand the test of time. For my purpose, the crucial core to all these proposals is the proposition that there exist specific identifiable mechanisms affecting the histories of languages continuously over stretches longer than a single generation. If this is true, which seems likely, then there must be some identifiable property of the language acquirer's experience which has the effect of inducing a competence different in some way from that of the previous generation. If such patterning in the input data were not possible, there could be no medium through which such long-term diachronic mechanisms could manifest themselves; the diachronic spiral through LAD and Arena of Use would not exist; languages would be only reinvented with each generation, and they would contain no 'growth marks', in the sense of Hurford (1987).

3. CONCLUSION

Language, in some broad sense, is equally an object of interest to biologists, to students of language acquisition, of grammatical competence, and of discourse and pragmatics, and to historical linguists. Each of these

disciplines has its own perspective on the object (e.g. focussing on E-language or I-language), but the perspectives must ultimately be mutually consistent and able to inform each other. The biological linguist is concerned with the innate human properties giving rise to the acquisition of uniformly structured systems across the species. The student of language acquisition is concerned with the interplay between these innate properties of the grammar representation system, other aspects of internal structure (e.g. innate processing mechanisms), and the learner's experience of the physical and social world. Students of discourse and pragmatics focus on, and hope to be able to explain and predict, certain patterning in the social linguistic intercourse which the learner experiences. Such patterning makes some impact on the grammatical competence acquired, resulting in the grammaticalisation of discourse processes, at which point the phenomena engage the attention of the student of competence. Frequency monitoring and individual creativity play a part in this diachronic spiral through grammars and use, by which languages develop, giving rise to the processes studied by the historical linguist.

The LAD is born into, and lives in, the Arena of Use. The Arena does not, in the short term, shape the Device, but, in conjunction with it, shapes the learner's acquired competence. The interaction between this competence and the enveloping Arena reconstructs the Arena in readiness for the entry of the next wave of LADs.

FOOTNOTES

1. Pinker and Bloom mention some of the evidence for this:

> "Bever, Carrithers, Cowart, and Townsend (1989) have extensive experimental data showing that right-handers with a family history of left-handedness show less reliance on syntactic analysis and more reliance on lexical association than do people without such a genetic background.
>
> Moreover, beyond the "normal" range there are documented genetically-transmitted syndromes of grammatical deficits. Lenneberg (1967) notes that specific language disability is a dominant partially sex-linked trait with almost complete penetrance (see also Ludlow and Cooper, 1983, for a literature review). More strikingly, Gopnik, 1989, has found a familial selective deficit in the use of morphological features (gender, number, tense, etc.) that acts as if it is controlled by a dominant gene". (Pinker and Bloom, 1990)

2. Sperber and Wilson's theory is, however, still controversial. See the peer review in *Behavioral and Brain Sciences* 10 (1987), also the exchange in *Journal of Semantics* 5 (1988), and Levinson (1989).

3. This is how Fodor (1976) casts a theory of language:

> "The fundamental question that a theory of language seeks to answer is: How is it possible for speakers and hearers to communicate by the production of acoustic wave forms?". (Fodor, 1976:103)

4. In this quotation, I have (with the author's approval) three times replaced an original instance of 'speakers' with 'language learners' and (indicating a shift in my opinion about certain numeral expressions) replaced 'preferred usage' with 'a fact of grammar'.
5. This convention is actually quite standard. Pinker, for example, adopts this usage: 'What the Uniqueness principle does is ensure that languages are generally not in proper inclusive relationships. When the child hears an irregular form and consequently drives out its productively generated counterpart, he or she is tacitly assuming that there exists a language that contains the irregular form and lacks the regular form, and a language that contains the regular form and lacks the irregular form, but no language that contains both". (Pinker, 1984:360)

REFERENCES

Andersen, H. 1973. Abductive and Deductive change. *Language* 40. 765-793.

Atkinson, M. 1982. *Explanations in the Study of Child Language Development*. Cambridge: Cambridge University Press.

Bates, E., I. Bretherton and L. Snider. 1988. *From First Words to Grammar: Individual Differences and Dissociable Mechanisms*. Cambridge: Cambridge University Press.

Bates, E. and B. MacWhinney. 1987. Competition, Variation, and Language Learning. In B. MacWhinney (ed.) *Mechanisms of Language Acquisition*. 157-193. Hillsdale, New Jersey: Erlbaum.

Benediktsson, H. 1963. Some Aspects of Nordic Umlaut and Breaking. *Language* 39. 409-431.

Beukema, F. and P. Coopmans. 1989. A Government-Binding Perspective on the Imperative in English. *Journal of Linguistics* 25. 417-436.

Bever, T. G. 1975. Functional Explanations Require Independently Motivated Functional Theories. In R. E. Grossman, L. James San and T. J. Vance (eds.) *Papers from the Parasession on Functionalism*. 580-609. Chicago: Chicago Linguistic Society.

Bever, T. G., and D. T. Langendoen. 1971. A Dynamic Model of the Evolution of Language. *Linguistic Inquiry* 2. 433-463.

Bever, T. G., C. Carrithers, W. Cowart and D. J. Townsend. (in press). Tales of two sites: The quasimodularity of language. In A. Galaburda (ed.) *Neurology and Language*. Cambridge, Massachusetts: MIT Press.

Bickerton, D. 1981. *Roots of Language*. Ann Arbor, Michigan: Karoma.

Borer, H. and K. Wexler. 1987. The Maturation of Syntax. In T. Roeper and E. Williams (eds.) *Parameter Setting*. 123-172. Dordrecht: Reidel.

Borer, H. and K. Wexler. 1988. The Maturation of Grammatical Principles. Ms. Department of Cognitive Science, University of California at Irvine.

Brown, C. H. 1984. *Language and Living Things: Uniformities in Folk Classification and Naming*. New Brunswick: Rutgers University Press.

Bybee, J. L., and M. A. Brewer. 1980. Explanation in Morphophonemics: Changes in Provençal and Spanish Preterite Forms. *Lingua* 52. 271-312.

Bybee, J. L. 1986. On the Nature of Grammatical Categories. *Proceedings of the Second Eastern States Conference on Linguistics*. 17-34. Ohio State University.

Cerdegren, H. and D. Sankoff. 1984. Variable Rules: Performance as a Statistical Reflection of Competence. *Langugae* 50. 333-355.

Chomsky, A. N. 1965. *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, A. N. 1979. *Language and Responsibility*. Hassocks, Sussex: Harvester Press.

Chomsky, A. N. 1986. *Knowledge of Language: its Nature, Origin, and Use*. New York: Praeger.

Chomsky, A. N. and H. Lasnik. 1977. Filters and Control. *Linguistic Inquiry* 8. 425-504.

Clark, H., and S. E. Haviland. 1974. Psychological Processes as Linguistic Explanation. In David Cohen (ed) *Explaining Linguistic Phenomena*. 91-124. Washington D. C.: Hemisphere Publishing Corporation.

Coker, C. H., N. Umeda and C. P. Browman. 1973. Automatic Synthesis from Ordinary English Text. *IEEE Transactions on Audio and Electroacoustics* AU-21. 293-8.

Coopmans, P. 1984. Surface Word-Order Typology and Universal Grammar. *Language* 60. 5-69.

Corbett, G. 1983. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croom Helm.

Dawkins, R. 1982. *The Extended Phenotype: the Gene as the Unit of Selection*. Oxford: Oxford University Press.

Downes, W. 1977. The Imperative and Pragmatics. *Journal of Linguistics* 13. 77-97.

Du Bois, J. W. 1980. Beyond Definiteness: The Trace of Identity in Discourse. In W. L. Chafe (ed.) *The Pear Stories: Cognitive Cultural and Linguistic Aspects of Narrative Production*. 207-274.

Du Bois, J. W. 1985. Competing Motivations. In John Haiman (ed.) *Iconicity in Syntax*. 343-365. Amsterdam: John Benjamins.

Du Bois, J. W. 1987. The Discourse Basis of Ergativity. *Language* 63. 805-855.

Edie, J. 1976. *Speaking and Meaning: the Phenomenology of Language*. Bloomington, Indiana: Indiana University Press.

Fidelholtz, J. L. 1975. Word Frequency and Vowel Reduction in English. *Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society*. 200-213.

Fodor, J. A. 1976. *The Language of Thought*. Hassocks, Sussex: Harvester Press.

Foley, W. A. and R. D. Van Valin Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Foley, W. A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Fries, C. C. and K. L. Pike. 1949. Coexistent Phonemic Systems. *Language* 25. 29-50.

Givon, T. 1979. *On Understanding Grammar*. New York: Academic Press.

Givon, T. 1986. Prototypes: Between Plato and Wittgenstein. In C. Craig (ed.) *Noun Classes and Categorization*. 77-102. Amsterdam: John Benjamins.

Gleitman, L. R., E. Newport and H. Gleitman. 1984. The Current State of the Motherese Hypothesis. *Journal of Child Language* 2. 43-81.

Gold, E. M. 1967. Language Identification in the Limit. *Information and Control* 10. 447-474.

Golinkoff, R. M. and L. Gordon. 1983. In the Beginning was the Word: a History of the Study of Language Acquisition. In R. M. Golinkoff (ed.) *The Transition from Prelinguistic to Linguistic Communication*. 1-25. Hillsdale, New Jersey: Lawrence Erlbaum.

Gopnik, M. 1989. A Featureless Grammar in a Dysphasic Child. Ms. Department of Linguistics, McGill University.

Greenberg, J. H. 1966. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (ed.) *Universals of Language*. Cambridge, Massachusetts: MIT Press.

Grimshaw, A. D. 1989. Infinitely Nested Chinese 'Black Boxes': Linguists and the Search for Universal (Innate) Grammar. *Behavioral and Brain Sciences* 12. 339-340.

Grimshaw, J. and S. Pinker. 1989. Positive and Negative Evidence in Language Acquisition. *Behavioral and Brain Sciences* 12. 341-342.

Gropen, J., S. Pinker, M. Hollander, R. Goldberg and R. Wilson. 1989. The Learnability and Acquisition of the Dative Alternation in English. *Language* 65. 203-257.

Hasher, L. and R. T. Zacks. 1984. Automatic Processing of Fundamental Information: the Case of Frequency of Occurrence. *American Psychologist* 39. 1372-1388.

Hawkins, J. A. 1990. A Parsing Theory of Word Order Universals. *Linguistic Inquiry* 21. 223-261.

Hock, H. H. 1976. Review article on Raimo Anttila 1972. *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan. *Languaqe* 52. 202-220.

Hooper, J. 1976. Word Frequency in Lexical Diffusion and the Source of Morphophonological Change. In W. M Christie, Jr. (ed.). *Current Progress in Historical Linguistics*. 95-105. Amsterdam: North Holland.

Horning, J. J. 1969. A Study of Grammatical Inference. Doctoral Dissertation, Stanford University.

Hudson, G. 1980. Automatic Alternations in Non-Transformational Phonology. *Language* 56. 94-125.

Hurford, J. R. 1987. *Language and Number*. Oxford: Basil Blackwell.

Hurford, J. R. 1989. Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device. *Lingua* 77. 245-280.

Hurford, J. R. 1991a. The Evolution of the Critical Period for Language Acquisition. *Cognition*.

Hurford, J. R. 1991b. An Approach to the Phylogeny of the Language Faculty. In J. A. Hawkins and M. Gell-Mann (eds.) *The Evolution of Human Languages*. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings vol. X. Addison Wesley.

Hyman, L. M. 1984. Form and Substance in Language Universals. In Brian Butterworth, B. Comrie and O. Dahl (eds.) *Explanations for Language Universals*. 67-85. Berlin: Mouton.

Ingram, D. 1979. Cross-linguistic Evidence on the Extent and Limit of Individual Variation in Phonological Development. *Proceedings of the 9th International Congress of Phonetic Sciences*. Institute of Phonetics. University of Copenhagen.

Itkonen, E. 1978. *Grammatical Theory and Metascience: a critical investigation into the methodological and philosophical foundations of 'autonomous' linguistics*. Amsterdam: John Benjamins.

Itkonen, T. 1977. Notes on the Acquisition of Phonology. English summary of: Huomiota lapsen äänteistön kehityksekä. *Virittäjä*. 279-308. (English summary 304-308).

Koopmans-van Beinum, F. J. and J. H. Harder. 1982/3. Word Classification, Word Frequency and Vowel Reduction. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 7. 61-9.

Kripke, S. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, Massachusetts: Harvard University Press.

Kroch, A. 1989. Language Learning and Language Change. *Behavioral and Brain Sciences* 12. 348-349.

Labov, W. 1969. Contraction, Deletion and Inherent Variability of the English Copula. *Language* 45. 716-762.

Lasnik, H. 1981. Learnability, Restrictiveness, and the Evaluation Metric. In C. L. Baker and J. J. McCarthy (eds.) *The Logical Problem of Language Acquisition*. 1-21. Cambridge, Massachusetts: MIT Press.

Lass, R. G. 1980. *On Explaining Language Change*. Cambridge: Cambridge University Press.

Lenneberg, E. H. 1967. *Biological Foundations of Language*. New York: John Wiley and

Levinson, S. C. 1989. A review of Relevance. *Journal of Linguistics* 25. 455-472.
Lightfoot, D. W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
Lightfoot, D. W. 1983. *The Language Lottery: Toward a Biology of Grammars*. Cambridge, Massachusetts: MIT Press.
Lightfoot, D. W. 1988. Creoles, Triggers and Universal Grammar. In C. Duncan-Rose and T. Vennemann (eds.) *On Language: Rhetorica, Phonologica, Syntactica: A Festschrift for R. P. Stockwell from his Friends and Colleagues*. 97-105. London: Routledge.
Lightfoot, D. W. 1989a. The Child's Trigger Experience: Degree-0 Learnability. *Behavioral and Brain Sciences* 12. 321-334.
Lightfoot, D. W. 1989b. Matching Parameters to Simple Triggers. *Behavioral and Brain Sciences* 12. 364-371.
Locke, J. L. 1986. Speech Perception and the Emergent Lexicon: an Ethological Approach. In P. Fletcher and M. Garman (eds.) *Language Acquisition: Studies in First Language Development* (2nd ed.). 240-250. Cambridge: Cambridge University Press.
Ludlow, C. L. and J. A. Cooper. 1983. Genetic Aspects of Speech and Language Disorders: Current status and future directions. In Ludlow, C. L. and J. A. Cooper (eds.) *Genetic Aspects of Speech and Language Disorders*. New York: Academic Press.
Macken, M. A. 1980. Aspects of the Acquisition of Stop Consonants. In Yeni-Komshian et al (eds.) *Child Phonology*. New York: Academic Press.
Macken, M. A. 1987. Representation, Rules, and Overgeneralization in Phonology. In B. MacWhinney (ed.) *Mechanisms of Language Acquisition*. 367-397. Hillsdale, New Jersey: Erlbaum.
McCawley, J. 1984. Review of White (1982). *Language* 60. 431-436.
MacWhinney, B. 1987a. The Competition Model. In B. MacWhinney (ed.) *Mechanisms of Language Acquisition*. 249-308. Hillsdale, New Jersey: Erlbaum.
MacWhinney, B. 1987b. Applying the Competition Model to Bilingualism. *Applied Psycholinguistics* 8. 315-327.
Mallinson, G. 1987. Review of B. Butterworth. B. Comrie and O. Dahl. (eds.) *Explanations for Language Universals*. Berlin: Mouton. *Australian Journal of Linguistics* 7. 144-150.
Martin, L. 1986. 'Eskimo Words for Snow': A Case Study in the Genesis and Decay of an Anthropological Example. *American Anthropologist* 88.2 (June). 418-423.
Martinet, A. 1961. *A Functional View of Language*. Oxford: Clarendon Press.
Miller, G. 1981. *Language and Speech*. San Francisco: Freeman.
Milroy, L. 1985. What a Performance! some Problems with the Competence-Performance Distinction. *Australian Journal of Linguistics* 5. 1-17.
Mithun, M. 1984. The Evolution of Noun Incorporation. *Language* 60. 847-894.
Moder, C. L. 1986. Productivity and Frequency in Morphological Classes. *Proceedings for the Second Eastern States Conference on Linguistics*. Columbus, Ohio: Ohio State University.
Mühlhäusler, P. and R. Harré. 1990. *Pronouns and People*. Oxford: Basil Blackwell.
Neu, H. 1980. Ranking of Constraints on /t, d/ deletion in American English. In W. Labor (ed.) *Locating Language in Time and Space*. 37-54. New York: Academic Press.
Newmeyer, F. J. 1980. *Linguistic Theory in America: The First Quarter-Century of Transformational Generative Grammar*. New York: Academic Press.
Newmeyer, F. J., forthcoming. Functional Explanations in Linguistics and the Origin of Language. *Language and Communication*.
Pateman, T. 1985. From Nativism to Sociolinguistics: Integrating a Theory of Language Growth with a Theory of Speech Practices. *Journal for the Theory of Social Behaviour* 15. 38-59.
Phillips, B. 1984. Word Frequency and the Actuation of Sound Change. *Language* 60. 320-342.

Piattelli-Palmarini, M. 1989. Evolution, Selection, and Cognition: From 'learning' to parameter setting in Biology and the Study of Language. *Cognition* 31. 1-44.
Pinker, S. 1984. *Language Learnability and Language Development.* Cambridge, Massachusetts: Harvard University Press.
Pinker, S. and P. Bloom. 1990. Natural Language and Natural Selection. *Behavioral and Brian Sciences* 13.
Pullum, G. 1989. The Great Eskimo Vocabulary Hoax. *Natural Language and Linguistic Theory* 7. 275-281.
Romaine, S. 1982. *Socio-Historical Linguistics: its Status and Methodology.* Cambridge: Cambridge University Press.
Saussure, F. de 1966. *Course in General Linguistics* (translated by Wade Baskin). New York: McGraw Hill.
Schane, S. A. 1971. The Phoneme Revisited. *Language* 47. 503-521.
Sperber, D. and D. Wilson. 1986. *Relevance: Communication and Cognition.* Oxford: Basil Blackwell.
Thiemann

Traugott, E. C. 1989. On the Rise of Epistemic Meanings in English: an Example of Subjectification in Semantic Change. *Language* 65. 31-55.
Wexler, K. and P. W. Culicover. 1980. *Formal Principles of Language Acquisition.* Cambridge, Massachusetts: MIT Press.
Wexler, K. 1981. Some Issues in the Theory of Learnability. In C. L. Baker and J. J. McCarthy (eds.) *The Logical Problem of Language Acquisition.* 30-52. Cambridge, Massachusetts: MIT Press.
White, L. 1982. *Grammatical Theory and Language Acquisition.* Dordrecht: Foris.
Wright, C. W. 1979. Duration Differences between Rare and Common Words and their Implications for the Interpretation of Word Frequency Effects. *Memory and Cognition* 7. 411-419.
Zipf, G. K. 1949. *Human Behavior and the Principle of Least Effort.* Cambridge, Massachusetts: Addison Wesley.