Robert Berwick
MIT

All you need is Merge:
Biology, computation, and language from the bottom-up

The Strong Minimalist Thesis (SMT) asks how little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles. It has recently been argued that interface conditions at the conceptual-intentional CI) side may largely fix the design properties of UG. In this talk we show that these design properties also dovetail nearly perfectly with constraints on the sensori-motor side, with, for example, the no-tampering condition (NTC), Edge labels, binary Merge, and the like all meshing with a computational model that imposes the minimal possible cost for SM. In this restricted sense, then, the entire system, from SM through to CI, is optimal.

Cedric Boeckx
Harvard University

Approaching parameters from below

In this talk, I'll examine the nature of parameters from a minimalist perspective. Given that parameters were embedded inside principles in the GB era, we expect the nature of parameters to vary once principles are reformulated in a minimalist way. I propose ways in which minimalist parameters may emerge from current work in syntax.

Alessandra Giorgi
Università Ca’Foscari di Venezia

Universal Grammar and temporal dependencies

In this talk I will consider temporal dependencies across languages from the point of view of Universal Grammar. Temporal relations are usually implemented in the various languages (mostly) by means of verbal morpho-syntax. Temporal, aspectual and modal morphemes codify the possible relations between events in main and embedded clauses. However, whereas on one side the inventory of the possible interpretive relations is universal — as well as relatively limited — on the other, verbal morpho-syntax varies a lot across languages. To exemplify, Italian and Romanian have a much richer morphology than English, and all of them have a richer morphology than Chinese — which has almost none. On the other hand, the ability of expressing a preceding relation, i.e., past or future, between events does not significantly differ among the languages I just listed and, in a way, we do not expect it to vary. The presence of very considerable morpho-syntactic differences is therefore rather puzzling. The important issue is to reduce language variability, by means of an appropriate level of abstraction, in order to predict and explain the arising of the various language types, by identifying the relevant dimensions and their universal properties.
Focus on biolinguistics — Focus in theoretical linguistics

This talk will address focus systems in grammar. On the basis of a discussion of the properties of the expression of focus in two closely related varieties of Greek (Standard Modern Greek and Cypriot Greek), coupled with a comparison to English, I will offer some preliminary remarks on the focus system from a biolinguistic perspective. More narrowly, I will also offer a novel analysis of cleft structures, this time applied to basic English *it*-clefts, extended to focus clefts in Cypriot Greek, and further adapted to *wh*-questions and more general predication structures. My proposal is intended to offer a contribution to both theoretical linguistics (by being couched in current minimalist theorizing) and biolinguistics as commonly understood (in this case, what possibilities grammar makes available and how they vary).

Biolinguistics: The “third factor” in evolution and variation

Biolinguistics studies language as a biological system and, as such, investigates language in the traditional areas of biology: form/function, development and evolution. The biolinguistic perspective concerns itself with the interplay between what Chomsky has termed the “three factors” in language design: 1) genetic endowment, 2) experience and 3) principles not specific to the faculty of language. We will consider the role of principles such as symmetry, which are non-domain specific and non-species specific, but which can contribute to the variation in biological systems seen in development and evolution.

On the limits of computation

I will explore some aspects of the question: 'What range of properties of the human language faculty can be ascribed to its computational component?"
Toward a history and geography of human syntax

Beyond its *theoretical* success, the development of molecular biology has brought about the possibility of extraordinary progress in the *historical* study of classification and distribution of different species and different human populations (e.g. Cavalli Sforza, Menozzi, and Piazza 1994). We want to suggest that even in the cognitive sciences purely theoretical progress in a certain discipline, such as linguistics, may have analogous historical impact, equally contributing to the so-called ‘New Synthesis’, and in turn be confirmed by such results. Thus, we will attempt to revive some classical concerns of the historical paradigm in linguistics at the explanatory level of the abstract entities provided by formal grammar and modern cognitive studies (cf. Guardiano & Longobardi 2005) and to unify two traditionally unrelated lines of investigation:

I) the formal study of syntactic variation (parameter theory) in the biolinguistic program
II) the reconstruction of relatedness among languages (phylogenetic taxonomy)

We claim, in particular, that, thanks to progress in grammatical theories and relying on the methodological parallelism with evolutionary genetics, we are now in the position of measuring the syntactic distance among different languages and populations in a precise fashion and to explore its historical significance.

Negation in the brain

Negation, consisting in reversing the truth value of affirmative sentences, is a unique human capacity: other animals can refute but not negate. What determines our ability to understand what is being negated? Using functional magnetic resonance imaging we measured brain activity in 18 healthy subjects during passive listening of sentences characterized by a factorial combination of polarity (affirmative vs. negative) and concreteness (action-related vs. abstract). Negation deactivated pallido-thalamo-cortical areas and, specifically for action-related sentences, reduced activity in the left mirror-neuron system. Thus, understanding negation may depend on the relative modulation of semantic representations.
The case of FOXP2 (the so-called language gene) revisited

Much publicity has been given to the identification of a specific mutation in the FOXP2 gene that causes various language impairments in the affected members of a British and Canadian family designated by the letters KE. Myrna Gopnik and collaborators had initially diagnosed a very specific "feature blindness" (problems with tense, number, agreement and gender) in otherwise perfectly normal subjects. A perfect case of dissociation between general intelligence and language. Subsequent refined neurological (PET and FMRI) analyses and a rich battery of tests have corrected this claim.

Orofacial dyspraxia and a limited capacity to learn complex sequences of oral movements have been proposed (contra Gopnik and colleagues) as the cause of reduced morpho-syntactic "skills" (sic). A reconstruction of the evolution of the Foxp2 gene in the mammalian world has prompted claims of an adaptive "sweep", from chimpanzees to humans, stressing the role of communication in the shaping of language. The most detailed genetic analysis today possible and a very detailed analysis of the neuronal correlates of the mutation seem to lead to a conception of language that may of us find quite problematic (language as part of motor planning and language acquisition as a chapter in the learning of complex motor sequences). I will present the case and suggest that this picture of language (offered by distinguished neurologists and geneticists, but non-linguists) is objectionable. A different lesson from this case, centrally involving the evolution of language, will be proposed.

On delimiting movement

Movement is a fundamental component of syntactic computations, with pervasive consequences on syntactic representations. Understanding the nature and properties of movement thus is a central component of the endeavour to understand the nature and properties of natural languages. So, the determination of the formal properties of movement has always been a crucial domain of syntactic research; more recently, the attempt has been made, within Minimalism, to connect the empirical discovery of the properties of movement with a deeper reflection on the nature and causes of the phenomenon.

A comprehensive formal theory of movement must include
1. locality principles, determining the maximal structural space which movement can cover;
2. delimiting principles, determining under what conditions movement can start, and must stop.

In this talk I will give a general overview of the issues, and then will focus on delimiting principles, with special reference to the cases which force a movement chain to stop and pass the representation on to the interpretive systems.
Juan Uriagereka  
University of Maryland

**Why knots, finches and more: Extending our data base**

One immediate consequence of embracing the ‘biolinguistics’ project is that the data-base of what constitutes evidence for the language faculty can be significantly expanded. The present talk provides two apparently unrelated examples (one from archeology and another one from neurobiology) which, when looked at a level of abstraction that only linguistic science can provide, end up significantly related, in terms of notions that use evidence from neurolinguistics and computational linguistics –in particular the role of memory in the well-known formal characteristics of grammars referred to as the Chomsky Hierarchy. The presupposed dynamic collegiality is central in the development of a new synthesis in linguistics, which stems from the Minimalist program seen from the Evo-Devo perspective in biology. This discipline is still crucially informed by the traditional findings of linguistics, but now these can be seen to affect subtle regulating mechanisms in development, and corresponding brain circuitry, particularly through the role of operational memory, whatever that turns out to be.

Charles Yang  
University of Pennsylvania

**The origin and diffusion of variations**

This paper presents a further step toward incorporating methods of evolutionary genetics into the study of language acquisition and change. We explore how, as a result of learning, syntactic variations may originate from and spread within a specific set of lexical items, and these variations then cause further changes in the global properties of the grammar. We show that the mathematical models of genetic drift can be used to capture the stochastic aspects of linguistic transmission, with implications in the study of linguistic typology as well.
Attributive modification and definiteness

In this paper we focus on the syntax (and semantics) of modification markers (MMs). MMs appear with adnominal modifiers. We restrict our presentation to adjectival modification. MMs appear in a variety of unrelated languages and are usually spelled out as (a) definite determiners (Greek Determiner Spreading/Polydefiniteness, Semitic definiteness agreement markers, Romanian and Aromanian ‘demonstratives’, ‘adjectival determiners’ in Albanian, Colloquial Slovenian, or Swiss German, Double Definiteness/Den-support in Scandinavian, for instance), (b) definite adjectival morphology (Slavic and Baltic) or (c) modificational particles (Chinese de or Persian ezafe). In the examples in p. 2 the MMs are in boldface. The contexts in (3)-(13) have been treated as being distinct phenomena. We argue for a unified analysis. From the study of the typology of MMs we concentrate on the following generalizations:

a) With the exception of Albanian, MMs do not occur with APs in predicative position.
b) There is a link between definiteness and the occurrence of MMs. In many languages, MMs occur only in definite DPs. In no language do MMs appear in indefinite DPs only.
c) MMs allow for a more flexible ordering of the adjectives with respect to the noun and each other; adjectives bearing an MM are more external with respect to the noun than bare adjectives (without MMs). Adjectival modification introduced by MMs can be characterized as indirect modification, cf. Sproat and Shih (1991), Cinque (2005).

Elaborating on Larson (1991 and subsequent work), we argue that determiners project DP shells which host the NP and the AP(s) and, in addition, that different definiteness domains (a definite and a non-definite one, for instance) can co-exist inside the DP. The link between definite DPs and the occurrence of MMs follows from the analysis of definite determiners as heads thematically selecting for a restrictive modifier, cf. Vendler (1967), Larson (1991).

Adapting Kayne’s (1997, 1998b, 2004) analysis of prepositional complementizers, we propose that an initial D-head is licensed on top of the thematic XP-complex in (1):

(1) \[ D_0 [\langle \text{XP} \rangle \text{NP} X_0 \text{AP}] \]

This \( D_0 \) attracts modifiers to a prenominal position and raises to the topmost D projection:

(2) \[ \underbrace{\langle \text{DP}_{\text{Max}} \rangle D_{0j}}_{\text{DP}^\text{Max}} \ldots \underbrace{\langle \text{DP}_1 \rangle \text{AP}_i t_j \underbrace{\langle \text{XP} \rangle \text{NP} X_0 t_i}}_{\text{XP}^\text{Max}} \]

Determiner Spreading phenomena and definiteness agreement phenomena are derived from these operations (or their iterative application). Language PF-independent conditions will be shown to determine the spell-out of the multiple occurrences of \( D_{0} \) in (2) (cf. Boškovic 2001).

Finally, we consider two aspects of the semantics of adjectival modification involving MMs: the relation between adjectives in predicative position and attributive adjectives involving MMs, and the relation between restrictiveness and MM adjectival modification. It has been claimed for certain languages featuring MMs that MM adjectival modification is restricted to intersective adjectives that can appear in predicative position (see, for instance, Alexiadou and Wilder 1998) and that MM adjectival modification is limited to restrictive modification (see, for instance, Kolliakou 2004). We argue that neither the predicative nature of the adjective nor restrictive modification are necessary conditions for MM adjectival modification.
(6) om-lu atsel bun-lu
    man-the that the good-the
    ‘the good man’

(7) marul (cel) ros
    apple-the CEL red
    ‘the red apple’

Aromanian

(8) ha-sepr ha-gadol
    the book the big
    ‘the big book’

Romanian

Hebrew


(9) d-i rot rosä
    the-AGRA red rose
    ‘the yellow shirt’

Swiss German

Norwegian

cf. Leu (2007)


(11) a. novijj bog?
    new-DEF god
    ‘the new god’

Old Church Slavonic

Chinese

Persian

cf. Lyons (1999)


References


Leu, Th. (2007) From Greek to Germanic: Poly-(*in)-definiteness and weak/strong adjectival inflection. Ms., NYU.


A case phase analysis of reconstruction and online binding

This paper argues that Binding properties of copies, as illustrated by Reconstruction and Online binding effects (“Barss” sentences”, 1986) derive from a Case phase analysis (Canac-Marquis 2005) according to which uninterpretable feature checking (e.g. Case) defines potential phase categories and subsumes local binding domains. The analysis aims to illustrate how core phenomena of language, such as the binding/reconstruction properties of copies, can result from the necessity of meeting the requirement of the C/I systems and the internal language at their interface.

Basic cases of Online binding are shown in (1) for A-movement and (2) for A-bar movement. The anaphor in (1), for instance, is located above the c-command domain of the antecedent, yet binding is licensed. Standard analyses for (1a) have maintained that either the subject constituent may reconstruct in object position to allow proper binding (Chomsky 1986, Barss 1986, Lebeaux 1988), or that anaphor binding is an “anywhere” condition (e.g. Belleti and Rizzi 1988), applying before or after A-movement. None of these proposals however seems compatible with a minimalist analysis in terms of derivation by phase (Chomsky 2000, 2001). For one, Lasnik (2003) argues convincingly that there is no reconstruction in A-chains. Further, if vP and CP are the relevant strong phases (Chomsky 2000;2001), it is technically impossible for the subject-raising constituent to spell out in vP -thereby allowing binding of its embedded anaphor at the C/I interface- before it reaches its surface position in TP where Case is checked. This follows from the fundamental minimalist assumption (Chomsky 1995) the uninterpretable features, such as Case, must delete prior to spell out to C/I interface. Related problems will be shown to apply to A-bar movement cases such as (2).

According to the alternative analysis of “online binding” developed here, phase domains correlate with Case checking projections, such as TP, AgrPhP, ApplP, PP. As Case must delete prior to spell out, categories involved in case-checking are the minimal phase or spell out points in a derivation. As shown in Canac-Marquis (2005), these minimal phase points actually correlate with local binding domains, as in (3). Basic binding examples such as (4-5) and others will be briefly reviewed. Under this proposal, the case-checking point of an XP also technically determines its lowest entry point at the C/I interface, i.e. the lowest reconstruction point of a XP. And this is supported as a fundamental generalization: reconstruction does not occur in A-chains (Lasnik 2003), as the head of the A-chain is case-checked, and reconstruction applies at tails of argument A-bar Chains, where case is checked.

As pointed out at the outset, Online binding such as (1) and (2) are challenges to a standard derivation by phase analysis. Under the Case phase approach though, Online binding phenomena can be can be traced back to the entry point of an anaphor at C/I interface. In (1), as the anaphor himself is case-marked by the preposition of, nothing prevents of himself to spell out as a member of the vP/AgrP phase prior to DP-raising to subject, as shown in (6). At that point, the experiencer object can also be case-checked and both DP enter C/I at the same vP/AgroP phase, satisfying Condition A in (3). Just as the spell out of himself can be anticipated in (1), it can in turn be delayed in (2): While the DP containing of himself has been case checked in the vP position, the whole DP is pied piped to check the Wh-feature and of himself can piggy bag on the DP until it reaches a point in the derivation (e.g. intermediate spec of CP position), where it can check its case and spell out, marking its entry point at C/I interface. That entry point corresponds to its local phase binding domain, which in this case is defined by the next case-checking category TP, which properly contains the antecedent: Condition A, as defined in (3), is satisfied. The paper discusses related cases, such as anti-reconstruction effects in (7a,b) and how reconstruction does not actually need to apply to allow binding in (8) and concludes on remarks to the effect that core properties of reconstruction/online binding phenomena are defined by the nature of the interaction of the internal language and the C/I systems at their interface.
1. a. [each other’s supporters]i frightened the candidates ti
b. [this picture of himselfi] seems to Johni to be horrible

2. a. *Johni thinks Mary likes [a picture of himselfi]?
b. [which picture of himselfi]i does Johnj think Mary likes ti

3. Binding condition A & B:
   Condition A: a reflexive is bound in its phase at spell out
   Condition B: a pronoun is free in its phase at spell out
   Legend ( = phase ; John = trace/copy ; John = phase Spell out
4. [TP[Johni] ([AgroP himselfi/*him [vP Johni likes himselfi /*him]])]
5. [TP[Johni] ([AgroP himselfi/*him [vP John believes [TP himselfi /*him to have won]]])
6. [TP ([AgroP [DP the candidate]i [AgrP[DP each otheri’s] supporters]j [vP frightened tj ti …]]

7. a. *[Which picture of Johni]j does hei like tj
b. *himi, Mary said that Johni likes ti

8. Which picture of himselfi]i does Johnj like ti

References
New thoughts on agrammatic Broca’s aphasia and the structure of the clause

Recent neurolinguistic research proposed that the syntactic tree is impaired in agrammatic Broca’s aphasia. Dissociations between TP (impaired) and AgrS (less impaired or even spared) led Friedmann & Grodzinsky (1997) to put forward a structural explanation for agrammatism, namely the Tree Pruning Hypothesis, which states that T is underspecified in agrammatic production and that an underspecified node cannot project any higher. Other approaches to agrammatism based on generative grammar have been formulated: Hagiwara (1995) claimed that cognitive limitations in the syntactic derivation (that is, in the operation of Merge) are responsible for several syntactic disorders. Other scholars argue that syntactic deficits do not rely on the operation of Merge, but involve the operation of Move (Bastiaanse & van Zonneveld 1998), Gavarrò (2002)). Finally, other studies on agrammatism have proposed a characterization of the deficit in terms of a failure in the feature checking operation (Wenzlaff & Clahsen 2004). The aim of this paper is to sketch a new linguistic description of some agrammatic disorders in Italian. We will consider data from sentence production and we will compare them with early reports in other languages. When comparative data turn out to be relevant for the discussion, I will present data drawn from bilingual aphasia (Italian- Northern Italian Dialects).

In this talk, I will assume a restrictive theory of movement (conceived as a computational operation taking place in narrow syntax) which bans head movement, following Koopman & Szabolcsi 2000. Secondly, following Manzini & Savoia (2005), I will assume that person features (or categories, in their own words) are interpretable (contra Chomsky 2000) and lexicalized in functional categories within the inflectional domain. Therefore, inflectional morphological processes take place within the syntactic component.

According to our data obtained through experimental tests on verbal Tense and Agreement, clitics, negation, interrogative phrases and adverb placement, we are going to assert that a structural approach - such as the ones based on the impairment of Merge-and-Move operations- cannot account for the (Italian/dialect speaking) non-fluent aphasic patients’ performance.

Furthermore, we are going to demonstrate that the syntactic tree, although impaired, is nevertheless present in agrammatic aphasia, and so are its functional projections.

We will also propose that both morphosyntactic and syntactic deficits in Broca’s aphasia may be satisfactorily explained in terms of movement deficiency, by assuming a restrictive theory of XP-movement (with its remnant part) which provides a more coherent explanation for those aphasic productions that previously used to be described as deficits concerning head-movement or maximal-projection movement.

References
Elements of bio-morphological variation

Considering the factors determining linguistic variation in the biolinguistic perspective (Chomsky 2004, 2007), I argue that morphological variation is biologically grounded. It is the result of the interaction of the genetic endowment with experience, and independent principles of efficient computation.

Because morphological variation, like syntactic variation, is biologically grounded, a unified understanding of linguistic diversity can be foreseen. I assume that the elements of bio-morphological variation are those that determine the growth of language in the individual (Chomsky 2004, 2007). I focus on the genetic endowment and the principles of efficient computation. Morphological merger and feature valuation derive constant hierarchical properties of morphological expressions. Fine-tuned morphological derivation by phase with relativized spell-out is a central element of bio-morphological variation. One consequence of the variation-by-phase approach to variation is that it both enables and imposes limits on possible morphological variation.

1. A(symmetry). Asymmetry breaks the symmetry of the natural laws, and brings about stability to an instable system (Jenkins 2000). While points of symmetry may arise in syntactic derivations (Moro 2000), they are never created in morphological derivations (Di Sciullo 2005a). The absence of points of symmetry in morphological derivations suggests that morphological derivations apply to elements, which have inherent asymmetric, stable, properties, which, according to our view are preserved through the derivations. The theory makes correct predictions for typologically different languages, as shown in Di Sciullo (2005b). Turkish is a strong suffixing language (Komfilt 2004, Sebüktekin 1971), nevertheless operator affixes precede the root. Yekhee, a North Central Edoid language from the Niger-Congo family, is a strong prefixing language (Bendor-Samuel 1988, Elugbe 1989), and operator affixes precede the roots while particles follow the root. The theory also makes correct predictions for Ethiopian Semitic languages, including, Chaha, Amharic, and Geez, as shown in Di Sciullo and Banksira (2007).

2. Phase and edge feature. I provide further evidence that Merge, which is part of the genetic endowment, thus a core element of FLN (Hauser, Chomsky, & Fitch 2002), applies in the derivation of morphological expressions. I also provide evidence that morphological derivations proceed by phases (Chomsky 2001, 2006, Kayne 2005, Boeckx & Grohmann 2004, Legate 2003, Uriagereka 1999). Both morphological and syntactic phases share the basic edge-head vs. complement asymmetry with respect to spell-out, while they differ with respect to edge features. Jointly, morphological merger, derivation by phases, and relativized spell-out yield the affixal precedence, recursion, and local dependencies.

The facts cannot be dealt with in terms of linear precedence relations, but follow from the properties of chunks of hierarchical structures recursively derived. From the biolinguistic perspective, this does not come as a surprise, since the language faculty is an aspect of the human biological nature, and is likely to be organized like the genetic code, with hierarchical, generative, and recursive properties.

3. Variation-by-phase. I discuss the variation in the pronunciation and silence of the preposition *a* ‘at’ in functional words denoting locations in Italian dialects, including Fallese, a dialect spoken in the Abruzzi (Catinella 1998, Di Sciullo 1996) and Piemontese. I argue that the variation between Fallese and Italian is due to a difference in phase spell-out (Chomsky 2006, Kayne 2005). Furthermore, I show that the pronunciation of the supplementary *a* in Fallese correlates with an independent morpho-phonological property of this dialect. Moreover, there is evidence that the supplementary *a* is a morphological remnant of Latin *ad*/*ab* PP structures, which suggests that historical change in the pronunciation/silence of a functional element can be reduced to differences in phase spell-out.

4. ERP. Morphological variation-by-phase is expected in the light of the experimental ERP results reported in Tsapkini, Jarema, and Di Sciullo (2004), which indicate that independently of morphological variation, humans process morphological constructs by phases. This also does not come as a surprise since phases contribute to reduce the processing load and are part of principles of computational efficiency, which are part of human biology.
References


Cedric Boeckx & Massimo Piattelli-Palmarini (in print) Language as a natural object, linguistics as a natural science. To be published in *The Linguistic Review*.


Chomsky, Noam. 2007. Approaching UG from below. Ms. MIT.


Di Sciullo, Anna Maria. 1996. Feature spreading in Fallese. Ms. Fallo and UQAM.


Neurocognitive research on language has mostly been dominated by two opposing views: the rule-based and the statistical learning theories. According to the former (Chomsky 1959, Pinker 1984, Marcus et al. 1996 etc.), the faculty of language, especially grammar, is an abstract and symbolic rule system, encoded in the human genetic endowment, allowing young human learners to bypass the induction problem when acquiring language and to use it with infinite creativity, once acquired. The latter view (Elman et al. 1996, Tomasello 2000 etc.), on the other hand, holds that language makes use of no dedicated mental and neural machinery. Rather it is acquired from the input in a piecemeal fashion as evidence accumulates.

These two learning mechanisms provide important insight into the architecture, and thus the possible evolutionary origins of language. However, for a fully-fledged biological understanding of language, at least one additional type of mechanism has to be taken into account: Gestalt-like primitive biases deriving from the architecture and functioning of the perceptual system. These ‘perceptual primitives’ have received extensive attention in vision research, but have mainly been neglected in the study of language.

In the present study, we report three experiments that explore such perceptual primitives in newborn infants, using near infrared spectroscopy. In Experiment 1, infants were presented with intermixed blocks of two artificial grammars. One of them generated trisyllabic sentences containing an immediate repetition (ABB: “mubaba”, “tofefe”, “pishosho” etc.), the other generated trisyllabic sentences without internal structure (ABC: “mubafe”, “tofesho”, “pishoge” etc.). We have obtained greater activation for the ABB grammar, mainly in the left frontal areas (Fig. 1), already in the initial blocks of the experiment, indicating that adjacent repetitions of the ABB kind are readily distinguished from stimuli with no obvious structure, especially in those brain areas that are responsible for structural representation building and integration (Dehaene-Lambertz et al. 2006). Even more interestingly, the response to ABB, but not to ABC, increased even further in the above mentioned brain areas during the 22 minutes of the experiment, providing a brain signature for auditory structural learning. These results imply that adjacent repetitions are immediately detected as perceptual primitives, and are later integrated into higher level, more general structural representations. In Experiment 2, we replaced the adjacent repetition ABB grammar with the non-adjacent repetition ABA grammar. In the case, no difference was found between ABA and ABA. In Experiment 3, we compared the ABC grammar to an A_A, where _ indicates a pause, teasing apart adjacency at the temporal and representational levels. Preliminary results suggest that the A_A grammar is distinguished from ABC one.

These findings argue for existence of auditory/linguistic perceptual primitives, such as representationally adjacent (but not non-adjacent) repetitions. These primitive, Gestalt-like configurations are automatically detected by the auditory system, and in later processing, they get recruited as building blocks of higher level representations.
Emergence of a systemic semantics through minimal and underspecified codes

In the pursuit of minimalist inquiries it is often assumed that semantics comes for free: semantics or ‘thought’ is what so-called ‘C-I’-systems incorporate, and the primary task is that of explaining the emergence of a syntactic system that interfaces with these pre-given systems, so as to express the relevant ‘thoughts’. Taking that as a basis, various syntactic constraints are rationalized as answering various ‘conditions imposed’ by the C-I-systems, in optimal ways. On this model, internal computational complexity matches some specific task domain (composing predicates, taking arguments, thematic relations, ‘discourse’, etc.). Here I further pursue an alternative, which I argue to be both conceptually superior and more in line with plausible principles of economy and conservativity in the building of adaptive complexity in biological evolution. In particular, I argue that syntactic structure should, in exact opposition to the standard minimalist line, be regarded as radically underspecified with regard to the semantic task performed; domain-general principles of organizing information economically lend themselves to semantic uses, they engender semantic consequences. The prevalent idea of ‘matching’ complexity contradicts the principle that biological evolution does not engineer novel algorithms for newly arising tasks. And it is not truly minimalist in duplicating semantic complexity with syntactic complexity where there is only one generative system based on minimal codes underspecified for the output it yields in different domains. A matching complexity model is preserved even in very recent work in the biolinguistic program that assumes semantic complexity to strictly track syntactic complexity (Uriagereka, forthcoming).

Hinzen (2006, 2007a) began pursuing an idea that abandons matching complexity through the assumption that structured thought emerges with the very computational system that provides these structures; moreover, I argued that the strongest hypothesis is that the computational system in question is (narrow) syntax, when appropriately minimalized (the ‘language of thought’ sensu Fodor (1975) is the syntax of thought). The externalization of these private languages or codes in a phonetic channel is a secondary and evolutionary independent development, irrelevant to the basic principles used for semantic purposes. As Chomsky (2006) paraphrases this architectural model of the faculty of language (FL), ‘optimally designed FL “provides forms that a possible human structured meaning may have”’, which he classifies as a ‘more radical conception of the FL-CI interface relation’. Strictly, it implies the ultimate non-existence of a ‘semantic interface’ and CI-systems, none of which, I argue, are ‘virtual conceptual necessities’. The present model abandons both for a ‘use theory of meaning’ in the sense of very early work in generative grammar.

Fleshing out this model further, I will focus on two issues in this talk. First, I argue that several longitudinal studies on seriation in monkeys (McGonigle and Chalmers, 2007) suggest the primacy of hierarchical over linear organization, which has long since been assumed for the case of language, even in the species investigated; they also suggest the apparently relatively free availability of non-domain specific, economic, and minimal relational codes that yield the famed semantic ‘systematicity’ claimed to be a prime feature of human linguistic thought in Fodor and Pylyshyn’s classic (1988). That is, systematic semantics begins where the relational codes and hierarchical units are in place, and it begins as a consequence of them, with their use for a linguistic purpose a historically contingent as well as gradual (as opposed to saltational) affair. What language adds to this basic hierarchical and relational organization are further principles of information management such as cycles of computation with exit strategies that then engender new cycles – all in the absence of invoking particular semantic task domains that motivate such cycles by expressive needs. A systematic semantics emerges on the basis of entirely domain-general principles whose minimality entails their underspecification.

The second issue dealt with is that if this picture is correct, another aspect of standard minimalism must be wrong, the strict uniformity of syntactic computation as recursive set-formation (e.g., Chomsky, 2006). If so conceived, as Hinzen (2007b) argues, structure-building will be non-hierarchical in the sense...
of never generating different categories: all it yields is infinite embedding of always the same category. Such a categorial non-productivity matches adjunct structures at best, and abandons the very principles of organizing information hierarchically and cyclically which lie at the heart of the origin of a systematic semantics in the animal world.

References
Jean-Rémy Hochmann, Mahan Azadpour, and Jacques Mehler
SISSA, Trieste

Do humans learn artificial center-embedded dependencies?

Showing that different animal species can learn grammars of different complexities is a fruitful approach to better understand the faculty of human language. Following Fitch and Hauser (2004, henceforth F&H), a series of studies assessed the ability of different species to learn two grammars, \((AB)^n\) and \(A^nB^n\). The latter, an example of a context-free grammar (Chomsky, 1963), is more complex than the former, a regular grammar. Indeed the \(A^nB^n\) grammar requires building non-adjacent dependencies, eliciting a center-embedded pattern. Acquiring an \((AB)^n\) grammar on the other hand, necessitates only learning local relations between the categories of constituents. Data suggest that humans, but not tamarin monkeys, learn an \((AB)^n\) grammar, whereas both learn a simpler \((AB)^n\) grammar (F&H). In a conditioning procedure, some songbirds, European starlings, can learn both grammars (Gentner et al., 2006). These studies, comparing different species is doubtless informative on the debate on language evolution. However, it is also crucial to assess what exactly is learned by humans in the original experiment of F&H and by the other species studied. In F&H experiments, the \(A\) constituents were syllables pronounced by a female voice, whereas the \(B\) constituents were syllables pronounced by a male voice. We propose that using such salient constituent categories may have distorted the results. In Experiment 1, we replicate F&H data. We then report new experiments using either similar categories as in F&H or less salient categories. In Experiments 2-4, we show that distributional regularities better explain the data than rule learning. Indeed, when familiarized with \(A^nB^n\) exemplars, participants failed to discriminate \(A^3B^2\) and \(A^2B^3\) from \(A^nB^n\) sentences. Therefore, contrary to F&H, we conclude that no syntactic rule implementing embedded non-adjacent dependencies were learned in these experiments. The difference between human linguistic abilities and their monkey equivalents or precursors remains to be experimentally identified.

References
Architectural constraints in the lexicon: The three-argument restriction

It is a universal property of argument structure that verbs cannot take more than three arguments – one external and two internal –, as in the English verbs give or put, unless if introduced either by an extra lexical preposition (1a), an applicative or causative morpheme (1b), or the additional head of a serial verb construction (1c) (Dixon and Aikhenvald 2000, Peterson 2007). In any case, the valency of the resulting predicate is restricted to that of a ditransitive verb. I call this the Two-Argument Restriction (TAR, see 1d), which remains one of the mysteries of human language. Since there is no known processing reason not to lexically associate more than three participants to a predicate, I claim that the TAR is syntactic in nature, and is one of a family of architectural constraints that determine and limit possible attainable languages (in this case possible argument structures). In this talk I aim to derive such generalization within the framework of lexical syntax put forth by Hale and Keyser 2002 (henceforth H&K). For H&K, the possible argument structures of verbs and their syntactic behavior are due in part to the structural properties of the basic lexical categories that compose them (see 2), mainly a head with a complement and no specifier, a head with a specifier and no complement, a head with a complement and a specifier, and a head with no complement and no specifier. (In many senses, this typology is equivalent to a syntactic one including unergatives, unaccusatives, transitives and impersonal verbs such as weather verbs, or one in which argument types are limited to S, A and O). All argument structures are obtained by the combination of these lexical categories. In H&K such combination is unrestricted in number (although syntactic restrictions apply), with the undesirable consequence that there is no limit on the number of arguments that can potentially be associated to a lexical head, thus violating the TAR. In my extension of H&K, the combination of basic lexical categories obeys the basic restriction that only one instance of each category is possible in any given composite resulting structure. I call this restriction the Uniqueness of Selection Hypothesis (USH, see 3). This amounts to negating the existence of a recursive function in the domain of argument structure, relegating such functions to the domain of sentences. I show that by adopting a simple restriction like that of the USH, the mystery of the Two-Argument Restriction can be solved. This is so because the USH both derives the most complex argument structure (that of a verb with two internal arguments and one external one), and accounts for the maximality of such structure: any larger structure would require recursive introduction of lexical categories, thus violating the USH. Moreover, it is speculated that the USH presumably derives from principles of the assignment of roots to abstract structures in a Distributed Morphology fashion.

References

(1) The Three Argument Restriction (TAR)

a. Erika gave Alexandra a present *(through/on behalf of) Frances

b. wa’-khe-yó:nye’ Meri ay-e-yo-h gwa’yea’
FACT-1sg.NOM.3ACC- make-PUNC Mary OPT-3.SG.NT.NOM-kill-PUNC rabbit
“I made Mary kill the rabbit”
(Baker 1996)
c. Olú ti omo nàà subú
   “Olu pushed the child down; Olu caused the child to fall by pushing him”

   (Baker 1996)

   d. “A single predicate must have at most two internal arguments and one external”

   (2) H&K’s Lexical Categories

<table>
<thead>
<tr>
<th>Monadic</th>
<th>Basic Dyadic</th>
<th>Composite Dyadic</th>
<th>Atomic</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+ cmp]</td>
<td>[+ cmp]</td>
<td>[- cmp]</td>
<td>[- cmp]</td>
</tr>
<tr>
<td>[- spc]</td>
<td>[+ spc]</td>
<td>[- spc]</td>
<td>[- spc]</td>
</tr>
</tbody>
</table>

   KEY: h = head; cmp = complement; spc = specifier; h* = non-projected head.

   (3) The Uniqueness of Selection Hypothesis (USH)

   An argument structure can contain two lexical categories x, y only if x ? y

   (4) Maximum Argument Structure according to the USH, predicting the TAR

   ![Diagram](attachment:diagram.png)
On determining the contents of Universal Grammar

The Tension Chomsky 2005 (building on Hauser et al 2002) emphasizes the tension between explanatory adequacy, which appears to require a rich Universal Grammar, and language evolution, whose understanding is facilitated by an impoverished UG. Chomsky states that in the ideal case, a single mutation would be sufficient, the operation Merge, and points out that “[t]he fundamental question of biology of language mentioned earlier then becomes, What else is specific to the faculty of language?” This is the question pursued in this paper.

A Problem Kiparsky 2004 proposes criteria, (1), to distinguish a typological generalization, which is explained historically/functionally, from a universal “which is properly the subject matter of UG.” These criteria result in a far richer UG than evolutionary considerations allow. We argue that they are misguided. To illustrate, we examine one of Kiparsky’s proposed universals, the nominal hierarchy underlying split ergative case marking (Silverstein 1976), (2), whereby nominals on the left of the hierarchy are more likely to exhibit accusative case marking, while nominals on the right are more likely to exhibit ergative case marking.

The Third Factor We argue that the nominal hierarchy is a “third factor” effect – one of the “principles not specific to the faculty of language” (Chomsky 2005), specifically a language-independent property of humans’ conception of the world, which manifests itself in various cognitive and behavioural domains, from the theory of mind to willingness to consume. Moreover, the manifestation in split ergative case-marking shows the same gradience found in other domains, e.g., just as we may anthropomorphize our pets, certain domesticated animals may “count” as human for case-marking purposes in many split-ergative languages.

An Analysis We propose a specific diachronic path whereby the non-linguistic nominal hierarchy leads to split ergativity in the synchronic grammar. Assume the beginning state, at generation $G_0$, to be a grammar in which ergative case and accusative case are both morphologically marked. Suppose that an innovation is introduced, whereby use of case morphology becomes optional. Optionality here can be understood as probabilistic accessing of the case morphology in the lexicon. The input for $I$-language acquisition of children at $G_{i+1}$ is the $E$-language of adults at $G_i$. Given the probability-matching effects in language acquisition, and the sensitivity to input frequency (Labov 1994, Yang 2002), the language acquisition device of the children of $G_i$ will assign a lower access probability to the morphological case markers which are less frequently heard in acquisition. Which morphological case markers will be less frequently heard? Here third-factor effects play a primary role, and the nominal hierarchy exerts its influence. As Silverstein (1976) noted nominals on the left of the hierarchy are felt to be more natural agents and those on the right more natural patients. Thus nominals in unnatural roles will be marked, while nominals in natural roles will be unmarked. Evidence for this claim is found in Japanese/Korean optional case omission, in which the nominal hierarchy plays a statistically significant role (Lee 2006). For split-ergative languages, the generationally-decreasing access probability for natural case markers, results in a generation for which the input data on natural case markers is insufficient for acquisition. The natural case markers are lost, resulting in obligatory split ergative case marking only on unnatural roles.

Outlook By factoring out of UG the universals that are due to general cognitive properties, we may approach the ideal situation for understanding language evolution.
(1) Universals (Kiparsky 2004)
   a. Irreversibility
   b. Convergence
   c. The Emergence of The Unmarked effects
   d. Manifested spontaneously in child language
   e. Pathways of change
   f. Structurally encoded in the grammar

(2) Nominal Hierarchy

<table>
<thead>
<tr>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>Proper/Kin</th>
<th>Human</th>
<th>Animate</th>
<th>Inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ERG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References
Kiparsky, Paul 2004. Universals constrain change; change results in typological generalizations. Stanford University
ms.
Rita Manzini and Leonardo Savoia
Università degli Studi di Firenze

(Bio)linguistic diversity

1. Chomsky (2004: 8) makes it explicit what the relation is between language variation and the faculty of language as a biological system. Variation can be seen as a correlate of ‘the growth of language in the individual’, in which the genetic endowment interacts with experience. The latter ‘leads to variation, within a fairly narrow range, as in the case of other subsystems of the human capacity and the organism generally’. On the basis of these conceptual premises we hardly expect variation to take the form of the broad generalizations entertained by the typological tradition, which takes as its starting point (not coincidentally) functionalist (hence fundamentally behaviorist) conceptions of language. Rather, the overall view that emerges from our study is microparametric, in the sense that parameters reduce to elementary properties of lexical items at the interfaces, and these generally combine freely, up to the general consistency of the system. If this conclusion is correct, it contributes an argument in favor of the biolinguistic perspective, to the extent that this is the only model capable of predicting such variation.

2. We shall consider one of the classical results of typological work, concerning the existence of an ‘animacy’ hierarchy (Silverstein 1976, DeLancey 1981, Dixon 1994), in particular ranking 1st and 2nd person higher than 3rd person. Perhaps the best known application of the hierarchy is in split ergativity phenomena. However the same split is pervasive in the deceptively familiar Romance languages. Perhaps its more macroscopic manifestation is in Central and Southern Italian dialects in which auxiliary selection is determined by the person of the verbal paradigm. The distribution more often reported in the literature (cf. Kayne 1993) has avere ‘to have’ in the 3rd person and essere ‘to be’ in the 1st and 2nd person, as in (1) and much variation appears to maintain is basic shape (avere in the whole plural, etc.). However careful investigation reveals that the reverse pattern is also found, with essere in the 3rd person and avere in the 1st and 2nd – though in (2) for instance this paradigm is limited to the singular, the plural presenting avere in all persons.

We argue that the key to the distribution in (1)-(2) – and what is more to the distribution in table (5) (Manzini-Savoia 2005, 2007) – is represented by the existence in Romance of dialects that allow for auxiliary avere (3) or auxiliary essere (4) over the entire paradigm. Since essere is the copula in all Italian dialects, in the absence of any further restrictions we expect it to embed participles on a par with any other nominal/ adjectival predicate – this immediately accounts for (4). To be more precise, while the copular interpretation appears to simply derive from the composition of the basic meanings of ‘to be’ and of the embedded nominal/ adjectival predicate, in the auxiliary - perfect participle, i.e. present perfect, reading some form of complex predicate formation (restructuring, reanalysis, etc.) takes place, whereby a single perfective state/event is interpreted.

The pattern in (3) is one step more complex. Evidently, languages like Calascibetta (English, etc.) have a constraint in place whereby the present perfect/ complex predicate interpretation requires the embedding of the perfect participle under transitive avere. We can speculate that complex predicate formation involves the identification of the EPP argument of the participle with the matrix EPP argument; languages like English or Calascibetta have an interpretive constraint in place whereby this requires the presence of a predicate with the transitivity/ external argument property.

Crucially, we argue that (1) and (1) simply represent the crossing of the essere vs. avere parameter independently exemplified in (3)-(4) with the animacy hierarchy independently needed to account for (among others) ergativity splits. In fact, we assume that what really is at stake is a split between discourse-anchored referents (1st and 2nd person, i.e. hearer and speaker) and eventanchored referents (3rd person i.e. absence of discourse-anchored hearer/ speaker properties). The crossing of the essere vs. avere auxiliary parameter in (3)-(4) with the 1st/2nd vs. 3rd person split could in principle lead us to expect two systems, i.e. essere in the 1st and 2nd person vs. avere in the 3rd – and avere in the 1st and 2nd person vs. essere in the 3rd, precisely as observed in (4) and (5).
3. Thus we conclude that parameters correspond to categorial (feature) splits, while variation depends on
the interplay of the various categorial splits (parameters) which is indeed free as long as it is consistent
with general grammatical principles.

(1)  *S.Benedetto del Tronto* (Marche)

\[ \text{s, } \text{f, a, } \text{fem, fer, a } \text{vnut, } \text{visto} \]

I.am you.are s/he.has we.are you.are they.have come/ seen

(2)  *Morcone* (Campania)

\[ \text{add}50, \text{a, } \varepsilon \text{ menuto/durmuto} \]

I.have, you.have, s/he is come/slept

(3)  *Calascibetta* (Sicily)

\[ \text{aju, a, } \text{amu, atu, anu } \text{vntu/ðrmuto} \]

I.have, you.have, s/he.has, we.have, you.have, they.have come/ slept

(4)  *Pescolanciano* (Abruzzi)

\[ \text{sng, si, } \varepsilon, \text{ sem, seto, su}[@(nn) } \text{manut/ðrmuto} \]

I.am you.are s/he.is, we.are, you.are, they.are come/ slept

(5)  

<table>
<thead>
<tr>
<th></th>
<th>1ps</th>
<th>2ps</th>
<th>3ps</th>
<th>1pp</th>
<th>2pp</th>
<th>3pp</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>Poggio Imp. Castelpetroso</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>A/E</td>
</tr>
<tr>
<td>(ii)</td>
<td>Roccasicura</td>
<td>A/E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(iii)</td>
<td>Capracotta</td>
<td>A</td>
<td>E</td>
<td>E</td>
<td>A/E</td>
<td>A/E</td>
</tr>
<tr>
<td>(iv)</td>
<td>Gallo Matese</td>
<td>A/E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(v)</td>
<td>Monteroduni</td>
<td>A/E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(vi)</td>
<td>Vastogirardi</td>
<td>A/A-E</td>
<td>E</td>
<td>A-E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(vii)</td>
<td>Colledimachine</td>
<td>E</td>
<td>E</td>
<td>A-E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(viii)</td>
<td>S.Benedetto T.</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(ix)</td>
<td>Vitticuso</td>
<td>A/E</td>
<td>E</td>
<td>A</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>(x)</td>
<td>S.Vittore</td>
<td>E</td>
<td>E</td>
<td>A/E</td>
<td>A/E</td>
<td>A/E</td>
</tr>
<tr>
<td>(xi)</td>
<td>Sassinoro</td>
<td>A/E</td>
<td>E</td>
<td>A/E</td>
<td>A/E</td>
<td>A/E</td>
</tr>
<tr>
<td>(xii)</td>
<td>Secinaro</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A/E</td>
<td>A/E</td>
</tr>
<tr>
<td>(xiii)</td>
<td>Guardiaregia</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xiv)</td>
<td>Agnone</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xv)</td>
<td>Bisceglie</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xvi)</td>
<td>Giovinazzo</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xvii)</td>
<td>Ruvo</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xviii)</td>
<td>Bitetto</td>
<td>E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xix)</td>
<td>Montenerodo</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xx)</td>
<td>Padula</td>
<td>A/E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>(xxi)</td>
<td>Castelvecchio S.</td>
<td>E</td>
<td>E</td>
<td>A-E</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>

The table above lists the linguistic features of various locations, with columns indicating different phonological or morphological characteristics (1ps, 2ps, 3ps, 1pp, 2pp, 3pp) and rows corresponding to different locations (i to vii). Each row represents a location with its linguistic features indicated. The features include various morphological and phonological marks, such as vowels and consonants, indicating the linguistic diversity across the locations listed.
Erika Marchetto and Luca L. Bonatti  
SISSA, Trieste

**Learning rules from an artificial speech stream: Early developmental differences**

Language acquisition requires both the ability to segment the speech stream into constituent words, and the ability to extract regularities underlying them. Much progress has been accomplished to study how infants can infer word boundaries on the basis of the statistical relationships between neighboring speech sounds (see Saffran, Aslin, & Newport, 1996), whereas the mechanisms involved in the extraction of structural information are still unclear (see Marcus, Vijayan, Rao, & Vishton, 1999).

Peña, Bonatti, Nespor, & Mehler (2002) proposed a dual mechanisms model, according to which two different computational processes might be engaged, one for the discovery of statistical patterns, and the other for the discovery of structure. They showed that adults could segment a continuous speech stream using a powerful statistical mechanism, but they could discover more complex structures only when they were exposed to a subliminally segmented stream. Here we investigate to what extent the same abilities are present in infants.

Participants in our experiments were familiarized with an artificial speech stream composed of trisyllabic nonsense words separated by short silences. In them, every A (first syllable) predicts exactly C (third syllable). The four words can be grouped in two “families”, defined as sets of words instantiating the same A-C dependency, with an intervening varying middle syllable. They were individually tested with a head-turn paradigm.

In Experiment 1 and 2, 13-month-old infants were familiarized with a segmented speech stream. They were tested with non-words, *i.e.*, sequences composed of three syllables appearing in the stream but arranged differently, and rule-words, *i.e.*, novel items with the same A and C position as words but with a different middle syllable, instantiating a long distance dependency. Infants looked longer to non-words, although both rule-words and non-words were novel items. This result suggests that infants recognized rule-words as “familiar” items, and they were able to extract its underlying structure.

In Experiments 3 and 4, 13-month-old infants were familiarized with a continuous speech stream. They were tested either with words and nonwords (Experiment 3), or with rulewords and nonwords (Experiment 4). Results showed that they could learn statistical patterns, but they failed to learn a structure. The failure in structure learning follows exactly Peña et al.’s (2002) results.

Experiments 5-6 investigated structure learning in 8 month-olds. Infants were exposed to the segmented speech stream, and they were tested with rulewords and nonwords. They failed to extract structural information. However, in Experiments 7 and 8, we tested infants with words and nonwords, and we showed that they could learn statistical patterns.

Together, these results suggest that the ability to identify structural information in a quasi-continuous speech stream follows a developmental trajectory. Only 13 month-olds could successfully learn the structure and the statistical patterns, thus they performed as predicted by Peña et al.’s (2002) model. On the contrary, 8-month-olds could only rely upon a powerful computation of statistical properties. Overall, these findings seem to support a theoretical model that might explain the early origins of vocabulary and grammar acquisition.

**References**


Children’s passives and the theory of grammar

In this paper we present work on the acquisition of Greek passives with the objectives: a) to evaluate the predominant theories that seek to account for late development of passives in children’s grammar, i.e., Borer & Wexler 1987, (B&W), Fox & Grodzinsky 1998, (F&G), and: b) to demonstrate how within Wexler’s (2004) recast of late development of passives in terms of phase theory (Chomsky 2000, 2001), data from children’s (interrogative) passives contribute to the understanding of the ban on A’-to-A movement in grammar. Hence, this work should be seen not only as an example of how syntactic theory assists in understanding (the stages of) language development, but, also, as how non-adult stages of language are able to shed light, in novel ways, to theoretical claims that require clarification.

Background

Terzi & Wexler (2002), (T&W), argue that the late acquisition of Greek passives results from children’s problems with A-chains (B&W, 1987), by contrast to F&G (1998) who claim that children perform poorly on verbal passives (of non-actional verbs in particular) as a side effect of problems they have with transmitting the external theta-role of the (active) sentence to the DP object of the preposition by in the passive. The latter approach was rejected in T&W (2002) since the Greek adjectival passives tested were exceptionally good, although they contained a by-phrase. There were two flaws in the above study however: a) adjectival passives are formed only with actional verbs in Greek, (1), hence, the (agent) theta-role of the active sentence does not differ from the theta-role assigned by by to its object DP in the passive. This fact alone is able to explain the good performance on adjectival passives (although the worse performance on their corresponding verbal remains puzzling). b) while actional, (2), and non-actional verbal passives, (3), as well as (actional) adjectival passives, (1), were tested with the by-phrase, none was tested without it.

Experiments, results, discussion

The present study aimed at filling the above gaps, in addition to confirming the results first reported in T&W (2002). We tested 40 children, age 3:06 to 6:06, divided in 4 age groups (the same 3 of the previous study, plus an older one). We also aimed at improving the methodology in Terzi & Wexler (2002), hence substituted two of their verbs with verbs forming better passives and easier to depict. Moreover, children had to choose from 3 rather than 2 pictures, presented simultaneously to them on a computer screen. The results confirmed T&W (2002), since adjectival passives were, again, better than verbal passives for all age groups (although in a less prominent fashion than before). Crucially, however, the present study revealed no significant difference between passives with and without the by-phrase. This result is of particular importance for non-actional passives, (5), since F&G (1998) predict them to be better without the by-phrase. Preliminary results on the acquisition of prepositions provide independent support, as it indicated that children may have not acquired the different properties of apo ‘by’ at these ages.

Wexler (2004) refines and updates B&W (1987), proposing that children perceive passive v_{def} as v*, hence do not allow movement (of the DP object) to Spec,T (nor can they move via Spec,v*, since no relevant features are present on it). Adults can move to passive Spec,T directly, as v_{def} is not a phase (but see Legate 2003). Adults also allow movement in wh-passives, (6). But, as Wexler (2004) notes, we do not know if adults perform improper movement in this case (i.e., from Spec,v_{def} to Spec,T) since they can move directly to Spec,T (and then to Spec,C), (7). On the other hand, children have to move to Spec,v_{def} since they perceive v_{def} P to be a phase, and we know they have acquired interrogative features (presumably also present in v_{def}). Thus, they perform Spec,v_{def} to Spec,T to Spec,C, (8), i.e., A’ to A (to A’) movement. If such a movement were allowed, children would be expected to form good wh-passives (recall that we know interrogatives to develop earlier than passives).

We tested the above 40 children on wh-passives as well, and found they performed either worse or non-different than in declarative passives, (9)-(10), demonstrating - in ways adult grammar could not - that movement from an A’ to and an A position is not allowed.
(1) a. O Janis ine sprogmenos (apo ti Maria).
   the John is pushed-ADJ by the Mary
   ‘John is pushed (by Mary).’

   b. *O Janis ine agapimenos (apo ti Maria).
   the John is loved-ADJ by the Mary
   ‘John is loved (by Mary).’

(3) O Janis sproxnete (apo ti Maria).
   the John push-pass by the Mary
   ‘John is pushed by Mary.’

(4) O Janis agapiete (apo ti Maria).
   the John love-pass by the Mary
   ‘John is loved by Mary.’

(5) Verbal passives of non-actional verbs

<table>
<thead>
<tr>
<th>Age: 3.06-4.0</th>
<th>Age: 4.01-4.10</th>
<th>Age: 4.11-5.08</th>
<th>Age: 5.09-6.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>w. by-phrase</td>
<td>0.45</td>
<td>0.35</td>
<td>0.45</td>
</tr>
<tr>
<td>wt. by-phrase</td>
<td>0.35</td>
<td>0.47</td>
<td>0.52</td>
</tr>
</tbody>
</table>

(6) Pjos sproxnete (apo ti Maria)?
   who push-pass by the Mary
   ‘Who is pushed (by Mary)?’

(7) Pjosi ti T [vdef sproxnete ti] Adults
   push-pass

(8) Pjosi ti T [ti vdef sproxnete ti] Children

(9) Verbal passives of actional verbs (with the by-phrase)

<table>
<thead>
<tr>
<th>Age: 3.06-4.0</th>
<th>Age: 4.01-4.10</th>
<th>Age: 4.11-5.08</th>
<th>Age: 5.09-6.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>0.68</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>Interrogative</td>
<td>0.25</td>
<td>0.53</td>
<td>0.45</td>
</tr>
</tbody>
</table>

(10) Verbal passives of non-actional verbs (with the by-phrase)

<table>
<thead>
<tr>
<th>Age: 3.06-4.0</th>
<th>Age: 4.01-4.10</th>
<th>Age: 4.11-5.08</th>
<th>Age: 5.09-6.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>0.45</td>
<td>0.35</td>
<td>0.45</td>
</tr>
<tr>
<td>Interrogative</td>
<td>0.37</td>
<td>0.47</td>
<td>0.32</td>
</tr>
</tbody>
</table>

References
Learnability of Recursion in Language

Since the seminal work of Hauser, Chomsky & Fitch 2002, the recursive property of language is taken to be the only uniquely human component of the faculty of language in the narrow sense (FLN), which distinguishes human from other species. Yet, this does not necessarily entail that the computational capacity of recursion is genetically endowed or “innately born with.”

In this paper, we report on our simulation results of an artificial neural network model, based on Self-Organizing Map (Kohonen 1990, 1995), that indicate the recursive property of language may be learned from positive data alone. These results imply that the computational mechanism of recursion may emerge from linguistic experience, which, together with the result of Fitch & Hauser 2004 that tamarins’ inability to acquire context-free grammars, suggests that the neurophysiological substrate for the computational mechanism of recursion may be a species-specific genetic endowment to human while the recursive computation itself may have to be learned.

We experimented on language identification task (Gold 1967) by Recurrent Self-Organizing Maps (RSOM) we designed. We trained RSOM with strings of category symbols of English sentences. Then, RSOM was subject to novel strings of various lengths with possible recursions in themselves. RSOM was able to identify, with surprising accuracy, grammatical strings and ungrammatical strings of random sequences of category symbols.

The architecture of RSOM is diagrammed in Figure 1. The competition layer $L$ and the context layer $M$ are both two-dimensional planes of $15 \times 15$ square units. An input string $x_i = [x_i,1, x_i,2, x_i,3, \ldots, x_i,j]$ is fed to the input layer, one symbol by one symbol, where $x_i,j$ is a category symbol of $k$-dimensional vector, 3-dimensional, encoded in 3-bit vector, in this study. For training and testing, grammatical strings are generated with an elementary set of context-free rewriting rules (1–7) for English, with category symbols, N, V, A, P, C, I, and D.

The weights $W_{Lu,v}$ between an input symbol $x_i,j$ and all the units on $L$, and the weights $W_{Lav}$ between all the units on $L$ and the context vectors $C$ of all the units on $M$, jointly determine the Best Matching Unit (BMU) on $L$. The BMU $L_{s,v}$ at the state $s$ will be the competition unit with the smallest sum of the Euclidean distances between the input vector and the weight $W_{Lu,v}$, and between $C$ and the weights $W_{Lav}$, each of which is depreciated by the apportion $\gamma$ and $(1 - \gamma)$, respectively, (8). The apportion $\gamma$ and $(1 - \gamma)$ fix the relative importance between the input and the context “knowledge” attained up to $s$, in determining $L_{s,v}$.

Once $L_{s,v}$ has been determined for an input $x_i,j$ at $s$, its weights and the weights of its neighboring units are updated toward the input vector and $C$. Then, $C$ is attenuated by the rate $\beta$, and a dividend of the input vector at the rate of $(1 - \beta)$ is added to the context unit $M_{s,v}$, on the same coordinates as $L_{s,v}$, in accord with (9, 10). This process is repeated for each category symbol, one by one. When an entire input string is fed, all the weights are adjusted and $M$ is flushed. Then, the next string is fed, one symbol by one symbol. When RSOM is trained with all the training strings, the next training cycle $T + 1$ begins. After the training is finished, all the weights are frozen, and all the training strings are fed one final time, determining BMUs for each symbol, without updating any weights. Instead, the system records the coordinates $u, v$ of the BMUs for each symbol on the BMU List.

In simulation, if all the BMUs of a test string are on the BMU List, we judged the string as grammatical, while if any of the BMUs, even one, of a test string is not on the BMU List, the string is judged ungrammatical. With these criteria, we trained RSOM with 1,000 grammatical strings, and tested with 20,000 novel grammatical strings and 20,000 strings of random sequence of symbols, both of variable lengths.

The result was that 98.73% of grammatical strings were judged correctly, and 86.43% of ungrammatical strings were judged correctly, 92.58% accuracy on average.
Fig. 1. A schematic RSOM architecture.

(1) CP $\rightarrow$ C IP
(2) IP $\rightarrow$ DP I $\left\{ \begin{array}{l} VP \\ AP \end{array} \right\}$ (PP)
(3) DP $\rightarrow$ (DP) D (NP)
(4) NP $\rightarrow$ A* N $\left\{ \begin{array}{l} PP \\ CP \end{array} \right\}$
(5) PP $\rightarrow$ P (DP)
(6) AP $\rightarrow$ A* A $\left\{ \begin{array}{l} PP \\ CP \end{array} \right\}$
(7) VP $\rightarrow$ A* V $\left\{ \begin{array}{l} PP \\ DP \\ CP \end{array} \right\}$
(8) $L_{v} = \arg \min_{L_{v}} \left\{ \gamma \| x_{i,j}(s) - W_{L_{v}}^{in}(s) \| + (1-\gamma) \| C(s-1) - W_{L_{v}}^{con}(s) \| \right\}$
(9) $C_{M_{v}}(s) = \beta \times C_{M_{v}}(s-1) + (1-\beta) \times x_{i,j}(s)$
(10) $C_{M_{v}}(s) = \beta \times C_{M_{v}}(s-1)$

References: