# biolinguistics

XP

ZP

...

γ

α

XP

# Linguistic Universals as Human Universals — Divergent Views and Converging Evidence on Language Congruence and Diversity

Dietmar Zaefferer  &  David Poeppel

## Introduction

Ping-pong is a game linguists are quite familiar with: "Experience […] leads to variation, within a fairly narrow range, as in the case of other subsystems of the human capacity and the organism generally" (Chomsky 2005: 6). For one school of thought, variation is a dwarf: Ping! — "The true picture is very different: [L]anguages differ so fundamentally from one another at every level of description (sound, grammar, lexicon, meaning) that it is very hard to find any single structural property they share" (Evans & Levinson 2009: 429). For other schools of thought, variation is a giant: Pong!

It is interesting to note that linguistics is not the only discipline that is frequented by discussions in line with the maxim: If you exaggerate your point, I overstate mine in the opposite direction. In the preface to his book about religious ideas, the French psychologist and anthropologist Pascal Boyer reports that his colleagues "all laughed heartily" when he told them about his project of a "general theory of religion, explained in terms of universal cognitive processes […] Indeed, in most academic institutions […] the project would have seemed crazy" (Boyer 1994: xv). And the anthropologist Donald Brown notes in his book on *Human Universals*: "Some anthropologists write about universals with little or no sense that they are controversial, but other anthropologists — some very prominent […] — maintain that universals have little significance if they exist at all" (Brown 1991: 54).

With this kind of problem in mind, a workshop was organized (initially by DZ, but with increasingly indispensable support from DP) around the theme 'Universals' at the 2008 annual meeting of the Deutsche Gesellschaft für Sprachwissenschaft (German Linguistic Society) in Bamberg. Barring any background information, a linguistically trained audience would approach such a meeting with relatively clear presuppositions, for example anticipating a focus on cross-

linguistic data and on discussions of putative counterexamples to putative universals. What set this session apart from other, more linguistically focused workshops on universals was the desire to emphasize the big picture and hence to be both thematically and methodologically much more ecumenical, a desire reflected already in the workshop title, *Foundations of Language Comparison: Human Universals as Constraints on Language Diversity* (http://www.itl.uni-muenchen.de/forschung/tagungen/human_universals/index.html). The idea of the meeting was to find out about the underpinnings of linguistic universals — definitional universals ('What makes the cluster of phenomena defined by the/a notion of language coherent?') as well as empirical ones ('Which non-definitional features cluster around the definitional properties and why?') — by using insights from both within and beyond the field of linguistics in order to determine the place of linguistic universals among the human universals in view of ultimately explaining the former in terms of (some of) the latter. As it turned out, there was also a desire, at that point left unstated, to examine or connect with the research program of biolinguistics, broadly construed.

Since human universals concern both the human body with its brain and mind and the cultures and societies humans grow up and live in, contributions were invited from the following fields (in alphabetical order): anthropology, biology, cognitive science, linguistics, neuroscience, and sociology. Although the sought-after diversity was not entirely reached, the composition of the group of speakers was laudably varied. Apart from general considerations (Dietmar Zaefferer), the workshop included presentations on syntax proper (Boban Arsenijević & Wolfram Hinzen, Ljiljana Progovac, Joana Rosselló, and Hedde Zeiljstra), research in anthropology and theory of evolution (Christoph Antweiler and Peter Richardson), and experimental approaches (Tom Bever, Rainer Dietrich, Adriana Hanulíková, Jeff Lidz, Asifa Majid, Andrew Nevins, David Poeppel, Friedemann Pulvermüller, and Michael Ullman).[1] It goes without saying that such diversity also provides particular challenges to the audience and speakers.

What did we learn? The lectures and discussions highlighted, in a productive and provocative manner, the heterogeneity of the aspects of the workshop topic. That is to say, the research questions and tentative answers surrounding universals research in general are not much more heterogeneous than the concepts and methods of the biolinguistic enterprise in special. This is not completely unexpected, but such a workshop (and a selection of papers from such a workshop) can sensitize researchers from the different domains to some of the considerations central to neighboring disciplines. The fact that this group of scholars attended such a workshop to begin with underscores their willingness to learn about the conceptual architecture of related research areas and to entertain the benefits and limitations of interdisciplinary research.

The question of universals continues to offer a fertile ground for debate regarding the radically different explanatory attempts to derive the systematicity so ubiquitous in the human language system. One salient intellectual position in this debate can be seen in the recent work of Evans & Levinson (2009) mentioned above, whose research agenda is motivated by the desire to explain language

---

[1]     For a short summary of all workshop presentations compare Zaefferer & Poeppel (2008).

universals from generic, nonspecific factors. Although they set out to show "how few and unprofound the universal characteristics of language are" (p. 429), they don't deny that "there are significant recurrent patterns in organization". However, their claim "that these are better explained as stable engineering solutions satisfying multiple design constraints, reflecting both cultural-historical factors and the constraints of human cognition" (p. 429) is at variance with core assumptions of the biolinguistic program, which represents a second, equally prominent position in the discussion.

Although the biolinguistic axiom that "nothing in language makes sense except in the context of the biology of grammar" (Boeckx & Grohmann 2007: 3) is not necessarily incompatible with explanations based on cultural-historical factors, given that culture and history need a biological basis to live on, such a view would trivialize the notion 'biolinguistics' (cf. Koster 2009: 92). According to a non-trivial interpretation of this notion, cultural-historical factors are not within the purview of the biology of grammar, and therefore their treatment on a par with the constraints of human cognition by Evans & Levinson must be viewed as a clear and important point of disagreement with biolinguistics.

A third source of hypotheses regarding universals can be seen as deriving from the experimental research programs that attempt to identify the psychological and neurobiological infrastructure that forms the basis for knowledge of language, acquisition, language comprehension and production, as well as the neurobiological implementation of language. This kind of research is neither committed to minimalist assumptions, nor does it exclude cultural-historical factors. One of the merits of the present volume, we submit, lies in the fact that a series of suggestions are made about the human language processing system that come directly from empirical research. As a cautionary remark we add that the universal invariance of psychological and neurobiological infrastructure is mostly taken for granted (as a laudable exception we mention studies reported by Lidz in this issue that were conducted in Mysore, India).

On the whole, the collection of arguments presented here puts forward at least three types of evidence for universals. First, cultural-historical explanations aiming at identifying universals that follow from very high level constraints which interact with very generic and high-level properties of the speaker/listener. Second, considerations from the perspective of contemporary biolinguistics, where a cognitive science based analysis points to a highly restricted set of representations and computations that underlie linguistic competence and performance. Third, empirically identified properties of the human mind/brain that are made visible through experimental research.

The assumption that linguistic universals are human universals, which served as a backdrop for the Bamberg call for papers, is certainly rather innocent and unassailable if it is taken alone and by itself. Far less trivial is the related question of the domains where the properties of human languages are most noticeably correlated with non-linguistic properties of human beings: the domain of more or less clearly identifiable biological organs, the domain of historical-cultural artifacts, or both to an equally strong degree.

Similarly for the claim that genetic endowment, experience, and language-independent principles contribute to language development in the individual

(Chomsky 2005): Undeniable as this might be, it is rather controversial (i) what exactly the three factors consist of, (ii) how they interact in the individual, and (iii) how this interaction is modulated by the respective environment, especially the shared distributed mind the individual comes to participate in (Zaefferer 2007).

Against this backdrop it is curious (although by no means surprising) to note that the Bamberg workshop offered considerably more ideas about the constraints that might explain the congruence of human language and cognition than about the determinants of linguistic and cultural-historical diversity. This seems to reflect quite truthfully the distribution of forces in the current language-related scientific fields: Only Richardson & Boyd outline an explanation of why languages are not less diverse than they are, all other papers contribute more or less directly to answering the opposite question of why languages are not more diverse than they are (and nothing to the question why they don't coincide).

Another notable reflection of the thematic priorities in the current debate is the fact that none of the contributions challenged or even addressed the conjecture formulated by one of the guest editors (DZ) at the Bamberg meeting that the boundaries provided by biological constraints leave a tremendously vast space for variation that is used only to a minimal extent by existing and possibly evolving languages, and that therefore many interesting constraints cannot be in principle explained by anything close to biolinguistics in the non-trivial sense because they are consequences of cultural universals.

Returning to the research agendas that are represented in the present volume, an informal taxonomy suggests four flavors. One approach derives from traditional research in generative linguistics, enriched with experimental data. The conclusions of Lidz, as well as those of Nevins, can be characterized as such. The former illustrates his point with three case studies from syntax learning (on the basis of Germanic, Romance, and Dravidian language data) and the latter reports experiments designed to test phonological universals using artificial gram-mars. Both claim that their findings can only be explained by assuming a highly constrained hypothesis space (aka Universal Grammar[2]) that biases the learner towards the observed behavior.

A second approach is advocated by researchers more closely aligned with the biolinguistic research program. The work of Arsenijević & Hinzen and that of Progovac can be seen in that light. The considerations made in that research tend to focus on work in syntax, and a critical hypothesis states, for example, that the key universal of the human linguistic system and a decisive step in its evolution is the operation of recursion, a recurring topic at the meeting. Interestingly, while conceding that recursion is present in language, Arsenijević & Hinzen claim that this is much less relevant than commonly assumed in that recursion in language is not truly causally efficacious. Drawing also on language evolution, Progovac proposes to take the early predecessors of modern syntax seriously and argues against treating Move as the default option. According to her proto-syntax

---

[2]    It might be interesting for typologists to note that Nevins conceives of Universal Grammar as "a set of analytic biases that prefer certain language types over others" in his contribution to this issue.

scenario, syntax began with small clauses and subordination lacked the option of move, making Subjacency (restrictions of Move) the default. In decomposing syntax into its underlying evolutionary components she also addresses Poeppel's warning against the possibility of cross-sterilization between linguistics and the neurosciences by proposing neurobiological correlates of syntax.

As a third strand of investigations, psychologically and neurobiologically motivated research such as that described by Bever & Poeppel as well as by Pulvermüller aims to identify universals based on experimentation. Bever & Poeppel summarize an older idea whose time has come (back), the heuristic algorithm originally called Analysis by Synthesis. It is a 'recipe' for how perception might be organized, both across languages and across levels of linguistic representation. Pulvermüller, in his work, summarizes the state of his own cognitive neuroscience research with respect to the language-related brain regions, the time course of linguistic processes, the multimodal action-perception circuits and the prewired structural information.

Finally, a rather different perspective on human language universals comes from the high level considerations of anthropology and evolutionary biology, as exemplified by Richardson & Boyd. Taking as starting point cross-species comparison, they outline a possible scenario of why language in its uniquely human form might have evolved at all that emphasizes gene-culture coevolution in order to account for its tightly interwoven cultural and biological aspects without concealing that there is still considerable debate about the details of the division of labor between genes and culture in this process. They furthermore trace back the uniqueness of language to the uniqueness of human cooperation and they offer, unique themselves in this among the contributors, an answer to the question why languages are not less diverse than they are: The evolutionary advantage of diversity lies in limiting communication between people who cannot freely trust each other's truthfulness or who even with truthful messages would cause maladaptive behavior on the part of listeners. In other, slightly paradoxical, words, according to Richerson & Boyd an essential (and probably mostly culture-driven) universal property of languages consists in their tendency towards diversity.

Looking back, we feel that the contributions from different fields that constitute this volume give a realistic picture of the state of the knowledge in this domain. There is still much to be improved in order to get beyond bold exaggerations, blunt polemics, and mutual ignorance between the language-related sciences, because the danger of cross-sterilization is real and overcoming interdisciplinary barriers requires hard work, which to the detriment of the field is often not given proper credit. Still, we think that the experience with our workshop and with the long process of thoroughly reviewing and carefully revising its written outcome justifies our hope that the prospect is real that all participating disciplines will contribute to making the field move slowly but persistently forward towards an increasingly complete picture of the ways language universals relate to other human universals and general human diversity relates to linguistic diversity.

## References

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The *Biolinguistics* manifesto. *Biolinguistics* 1, 1–8.

Boyer, Pascal. 1994. *The Naturalness of Religious Ideas. A Cognitive Theory of Religion.* Berkeley, CA: University of California Press.

Brown, Donald E. 1991. *Human Universals*. New York: McGraw-Hill.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32, 429–492.

Koster, Jan. 2009. Ceaseless, unpredictable creativity: Language as technology. *Biolinguistics* 3, 61–92.

Zaefferer, Dietmar. 2007. Language as mind sharing device: Mental and linguistic concepts in a general ontology of everday life. In Andrea Schalley & Dietmar Zaefferer (eds.), *Ontolinguistics: How Ontological Status Shapes the Linguistic Coding of Concepts*. Berlin: de Gruyter, 189–222.

Zaefferer, Dietmar & David Poeppel. 2008. Bericht der Arbeitsgruppe 2 "Foundations of Language Comparison: Human Universals as Constraints on Language Diversity" auf der 30. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft in Bamberg, 26.–28. Februar 2008. *Mitteilungen der DGfS* 67, 36–39.

*Dietmar Zaefferer*                        *David Poeppel*
*Ludwig-Maximilians-Universität München*   *New York University*
*Institut für Theoretische Linguistik*     *Department of Psychology*
*Schellingstr. 7*                          *6 Washington Place*
*80799 München*                            *New York, NY 10003*
*Germany*                                  *USA*
*zaefferer@lmu.de*                         *david.poeppel@nyu.edu*

# Recursion as a Human Universal
# and as a Primitive

### Boban Arsenijević  &  Wolfram Hinzen

This contribution asks, in an empirical rather than formal perspective, whether a range of descriptive phenomena in grammar usually characterized in terms of 'recursion' actually exhibit recursion. It is concluded that empirical evidence does not support this customary assumption. Language, while formally recursive, need not be recursive in the underlying generative mechanisms of its grammar. Hence, while recursion may well be one of the hallmarks of human nature, grammar may not be the cognitive domain where it is found. Arguments for this claim are briefly exposed and then discussed with respect to a selection of talks from the DGfS workshop on *Foundations of Language Comparison: Human Universals as Constraints on Language Diversity* that led to this special issue.

*Keywords:*    cyclic spell-out; interfaces; narrow syntax; recursion

## 1.    Introduction

Much recent discussion around language evolution has focused on recursion as a putatively universal design feature of language, in such a way that claims that some languages do not exhibit this feature have proved highly controversial (Everett 2005). According to Everett, cultural considerations enter into the determination of whether recursion is present in a language, resulting in a return to early 20th century claims about the culture-relativity of human nature. Dietmar Zaefferer, too, at the workshop, underlined the 'dual' biological and cultural underpinnings of modern *homo sapiens*, and limitations of the extent to which the current 'biolinguistic' program is comprehensive enough to answer all core questions about human language. Against Everett, and despite Zaefferer's cautionary words, we maintain a biolinguistic approach to recursion as a human universal, though 'deconstructing' it into a number of independent and more primitive factors that we argue underlie it.

The claim that I-languages are recursive devices has been one of the hallmarks of generative theory and the basis for its implementation of the Humboldtian dictum that language makes 'infinite use of finite means': The recursive devices are the finite means in question. With recursive rules in place, a grammar has the power to generate a potential infinity of sentences, when, as Chomsky

put it over 50 years ago, otherwise it would be "prohibitively complex" (Chomsky 1956: 116). This employment of recursion in generative grammar was part of a more general inspiration of the theory of grammar by recursive function theory. Recursive definitions were a useful device that simplified an analytical framework (Tomalin 2007). Later, the notion of recursion acquired biological and cognitive connotations, and the adoption of an intensional perspective was emphasized, according to which not the result of a computation matters, but how in fact it happens. The proper object of linguistic investigation, on this cognitive view, is the one specific way (of potentially infinitely many) in which the mind/brain generates the expressions of a language. The claim that language is recursive is now the empirical claim that this particular algorithm is a recursive function (reflecting standard observation in linguistics that a linguistic expression of a particular syntactic, semantic or phonological category may become part of another one, of the same category). That it is has been a central claim in generative grammar to this day. In Minimalism, in particular, recursion became encapsulated in the definition of the basic combinatorial operation Merge, which moreover was identified as the prime and potentially single biological innovation in the evolution of language (Hauser *et al.* 2002): A universal and language-specific element of human cognition as well as a primitive, not reducible to anything else. This sense of recursion as instantiated in Merge, where the value of a function is added to its domain, is the one under discussion in the present paper.

In our presentation, we claimed that:

1.  Language uncontroversially displays recursive capacities: It can generate an infinite sequence of embedded expressions of the same category. But what gives rise to this recursivity is the confluence of a number of different factors in language design, all having to do with the interfaces that the computational system of grammar forms with other linguistic and certain extra-linguistic systems (e.g., the discourse representation, the lexicon, the planning capacity). There is no clear empirical evidence of recursivity of the structure-building mechanism of grammar in isolation or autonomously from these interfaces, notably the syntax-discourse interface (hence, there is no recursion in the faculty of language in the 'narrow' sense of Hauser *et al.* 2002).

2.  Considering syntax to be the module, or aspect, of grammar that drives the structure-building processes in language, 'direct' recursion — the immediate embedding of one and the same syntactic category in itself — never occurs within the structure-building grammatical computations, which speaks in favor of a templatic view of this process (similar to views of language in certain types of construction grammar). As elaborated in more detail in the next section, we consider as direct recursion only those cases where unmediated embedding of one syntactic category in itself is attested, as well as ones where other categories interfere between the one occurrence of the category in question and the other but still all takes place within a single cycle of syntactic computation.

3.      Recursive structures as they arise from interface effects derive from the more primitive properties: The cyclicity of derivations, the categoriality of syntactic constituents, and the way reference to discourse entities is determined at cyclic boundaries.

We do not deny that potentially infinite embedding in language, illustrated in (1), is for real. Rather, we argue that it only emerges as the consequence of a conspiracy of syntactic structure-building and the interfaces, and that it cannot be achieved by the former alone.

(1)     a.      [I saw the woman [that saw the woman [that saw the woman…]]]
        b.      [[The window [[the neighbor [the dog bit]] broke]] fell down]
        c.      [John knew [that Peter believed [that Mary liked him]]]

We argued that approaches to grammar assuming some version of Multiple Spell-Out — for instance, the phase-theory of Chomsky (2001, 2008b) — are forced to represent the narrow structure-building syntactic procedures as essentially non-recursive. In this view, every one of the embedded constituents in (1) is spelled out when completed and then interpreted at different interfaces, involving further processes outside the narrow syntactic derivation. Only a truncated constituent is included in any further structure-building operations in narrow syntax. As Chomsky (2008a, 2008b) argues, after Spell-Out, each constituent is structurally reduced to a structural primitive (comprising the head and the left edge of the spelled-out phase) which does not carry along any information about its complement. This means that any category embedded within a spelled-out chunk of structure becomes inaccessible for further computations, except for interpretively irrelevant processes of Agreement, as Chomsky (2001: 14) points out.

We presented arguments that direct recursion can never be observed in the operations constituting the structure-building (i.e. narrow syntactic) component of the language faculty, neither within nor between phases in cyclic computations. We discussed a number of different empirical regularities, from sequence of tense phenomena to complement clauses, in support of the argument that narrow syntax is non-recursive, in fact bans recursion. We pointed out facts as in (2), where embedded expressions of the same category show certain deficiencies: Clauses lack truth values, nominal expressions contribute descriptions, not immediate reference, and tenses lose the capacity to determine their own reference times.

(2)     a.      C–in–C
                [John suspected   [that Mary believed  [that he was a police agent]]]
                *truth value*          *no truth value*          *no truth value*
        b.      D–in–D
                The vase on the table was green.
                John's mother plays basketball.
        c.      T–in–T
                John said Bill was tired.

Possible counterarguments, such as the seemingly unrestricted compounding in the nominal domain, are shown to obey the same restrictions: Direct recursion must be avoided by a Spell-Out to the discourse and other interfacing domains, a special kind of this intermediate step being the storing of a compound in the lexicon before further compounding may take place.

We argued in the light of such facts that a templatic view of the structure-building component is better than the standard one based on the operation Merge. While the templatic nature of the phase-internal structure is not far from some of the prominent views of grammar within the minimalist community (especially in the so-called cartographic approaches, stemming from Cinque 1999), we argue that at the level of embedding of phases in one another, the same kind of templatic patterns plays a central role as well.

Section 2 discusses the relation between the contents of our talk and the contents of the talks taking the perspective of theoretical syntax. In section 3, we discuss how the papers in the domain of cognitive neuroscience relate to our views of recursion in language. Section 4 concludes.

## 2.      Grammar Architecture and the Core Properties of FLN

Our conclusion is that recursion is not an element of UG or of the 'faculty of language in the narrow sense' (Hauser *et al.* 2002), that is, not an element that is specific to grammar or characterizing language in some essential sense. This conclusion can be abstractly compared to a conclusion that Bever (2009) has suggested regarding such putative UG-principles as the 'EPP'. Rather than being a core universal constraint of language design, Bever argues that it reflects a non-domain specific constraint on learnability. Put differently, the EPP is an epi-phenomenon of learning: It is merely a descriptive universal, the overall result of recurrent statistical patterns in linguistic data that find its true causes in constraints on acquisition. In the words of Bever, the EPP is a property of "the connection between the narrow faculty of language and the acquisition interface" (p. 280). In an analogous way, we say that recursion is truly a property of the interface between the narrow faculty of language and the discourse in which language use takes place. Specifically, the EPP is the result of a 'Canonical Form Constraint' for Bever, which makes the learner identify statistically frequent templates in the linguistic input, which are then internalized and become a part of a speaker's I-language. This stance interconnects with our stance regarding the foundational significance of syntactic templates as opposed to the operations of unrestricted Merge in language.

While we argued against the central role of recursion in the narrow language faculty (FLN) by showing that the recursive nature of computations in language is epiphenomenal — it comes from its interface with external systems (such as discourse representation) — Joanna Roselló argued that the duality of the architecture of language, its double articulation, is a more essential property of FLN than recursion itself. Her main line of argument relies on the fact that recursive computations are found in other capacities, such as music or arithmetic, and that it is very hard to eliminate the possibility that these capacities are to

some extent independent of language. On the other hand, these two capacities may be difficult to motivate in terms of selective pressure, and also be parasitic on language (e.g., Hinzen 2008). In that case, however, we need to consider that capacities like planning or spatial cognition, which are both older than language and more broadly distributed in the animal world, are argued to include recursive computations (Arsenijević 2008, van Lambalgen 2008, and note especially the Sapir-Whorfian close relation between syntax and planning argued for at the workshop by Rainer Dietrich, Werner Sommer & Chung Shan Kao). This makes it quite hard indeed to see recursion as the core property of FLN. The final bearing of Roselló's talk on Hauser *et al.* (2002)'s view of recursion as the core of FLN is essentially the same: We argue that this view is untenable — for us because recursion in language is epiphenomenal and dependent on other modules instead of being internal to FLN, and for Roselló because there is a better candidate, which is an exclusive property of FLN, unlike recursion.

With Roselló, we also share the internalist view of language and the way it relies on computational capacities, as well as the approach to syntactic computations, in which they are taken as an instantiation of the general structure-building capacity of human cognition, hence as closely related to propositional thinking as they are to language. However, we have a slight reserve towards Roselló's view which puts the double patterning of language in the centre of FLN. There are cognitive capacities that we share with animals, which also involve the kind of duality of patterning that language involves. Such is the case, for example, with the planning capacity, where structures of plans always map between actions (sensory-motoric, just like the phonetic side of language) and goals (abstract representational, just like the semantic, or discourse side of language). Similarly, in the spatial capacity, there is the sensory-motoric and the cognitive map component.

Andrew Nevins discussed the possibility that there is a language that for cultural or other reasons lacks any recursive structures. Although this is highly unlikely, given that FLN is typically described as essentially recursive, he argued that this is still possible, and that it should be taken as not more than a rare accident. From the perspective of our view of recursion, as an epiphenomenon arising from the nature of the syntax-discourse interface, such an accident is not particularly unlikely — it only requires a certain version of that interface, which will block the patterns of interaction leading to the generation of surface — or formally recursive structures.

Another point of Nevins' talk was that even in languages where syntactic recursion seems to be missing, recursive computations are needed to deal with pronouns and paratactic expressions whose the semantic interpretations would involve embedding structures. In other words, even when recursion is expelled from the sentence, it is present in the discourse. This again fits our argument, insofar as we exactly claim that recursion in language comes from the interaction of syntax with the discourse.

Ljiljana Progovac presented a view of small clauses in which this class is at the root of not only syntactic derivations in a synchronic view of grammar, but also at the root of the evolution of syntax as we have it today. She presents small clauses as the first step in the derivation of any sentence, and as the first real

syntactic construction ever. Moreover, she shows that this primitive syntactic structure is still generated and used by humans, and that there are certain pragmatic and semantic domains in which it is still the first choice.

(3)   a.      John considers [her happy].
      b.      Her happy?!
      c.      *John considered [Bill see [her fall]].

Progovac discusses a number of special properties of small clauses, among which that they cannot be produced recursively, as illustrated by (3c).

An important question in this respect is that of the discourse-integration of bare small clauses, as in (3b). Are they treated as fully specified, that is, as expressions with a particular type of interpretation, determined by the structure and the content of the very expression, or are they treated as underspecified expressions, bearing a number of unspecified features, which get their specification in the discourse? The latter seems to be the correct view. Irrespective of the expressive content of the small clause in (3b), it is interpreted as expressing some attitude of the speaker towards the referent of 'her' in the relevant discourse being considered by the relevant subjects to be happy *at the time specified by the discourse/context*. This means that in the discourse, this expression binds a referent for its subject, and also gets a tense in a similar way (considering that tense behaves similar to pronouns). In the discourse, therefore, it behaves as a tensed expression.

From the aspect of our view of recursion, where recursivity emerges when a structure is taken from the discourse and used as an atomic (i.e. non-structurally complex or phrasal) element in syntactic generation, it is natural that small clauses do not recursively embed. Once integrated in the discourse, they receive tense and other specifications, and are not present as small clauses any more — especially considering that in our view only full phases can be turned into atomic elements that can form an input from the discourse to a new cycle of generation — a unit that presents only a root of a phase can never be taken as such an input.

## 3.      Neuro-Cognitive Aspects

The talk 'Linguistics and the future of the neurosciences' by David Poeppel discussed the still quite long distance between these two disciplines, stressing in particular the problem of different granularities: While linguistics works on fine-grained distinctions among different representations and computations, neurosciences use rather broad-stroke conceptual distinctions to characterize linguistic phenomena, and has its own lower-level vocabulary, possibly incommensurable with the linguistic one. The first step in solving this and related problems would be a sufficiently formal and abstract systematization of the ontologies and processes involved in grammar, which would allow for the modelling of neural populations responsible for their memorization and the execution of these ontologies and processes respectively.

Michael Ullman argued that there is a competition between different ways

of handling certain linguistic operations, and that although this involves a degree of redundancy, it also brings in considerable advantages in handling tasks of different kinds. In particular, Ullman argued that two distinct systems, one based on lexical memorizing, and one on productive generations, are engaged in the computation of expressions with morphologically or lexically marked functional features such as tense or definiteness.

In our interpretation, this view implies that the neurocognitive reality of grammar is somewhere between the more fully generative models such as the Minimalist Program (Chomsky 1995) and Construction Grammar (Goldberg 1995) in which there is a significant role of memorization of larger structures. Moreover, Ullman reports on evidence of a more frequent utilization of processing strategies compared to memorization in children, and in turn a more prominent role of memorization in adult speakers. This is in agreement with the option we discussed in our presentation, that even relatively large cartographic structures, specifying the full projection capacity of a certain category, are memorized by adult speakers, and then used as templates, but that they all still need to be generated a sufficient number of times before they are memorized due to their frequency.[1] This pre-memorizing productive generation takes place in the L1 acquisition period. This hopefully presents a step towards more commensurability between linguistics and neurosciences, in respect of Poeppel's concern.

Although this rule-based view is often discarded by neuroscientists taking network approaches, on the grounds that they are too discrete in nature to appear as a product of neural activities, Friedemann Pulvermüller presented a possible neuronal model capable of representing and executing discrete rule-based operations such as those typically defined by phonologists, syntacticians, and semanticists. He also presented a possible way of handling recursive rules within network approaches in neuroscience, an important requirement being the sensitivity of the system to the intensity of activation, which is somewhat related to our claim that recursion appears only once the structure-building capacity has reached the interface with the discourse, stored its output there, and read a new package of input from it. The execution of a procedure matching a certain operation of syntax or phonology in itself cannot produce a recursive structure; such a structure can be generated only once the system is made sensitive to the outputs of earlier executions.


## 4.    Conclusion

Although coming from very different domains of inquiry into the nature of language, presentations at the workshop converged on a large number of questions. In some cases, these questions are left open, or different answers were advocated for them in different talks, but there was also a great degree of convergence on the central issues of the conference topic: For example, the nature

---

[1]    Note that these are in fact not particularly large structures, but rather ordered linear sequences of 10 to 15 category labels. They may be thought of in terms of, for example, Gärdenfors' (2000) conceptual dimensions, which he also presents as ordered linear structures.

of recursive computations was one of the central topics in most of the talks, and it was approached in different ways and from different perspectives. If we are right, recursivity is present in language but not truly causally efficacious and not language-specific: Language-specific are specific restrictions on what ultimately templatic structures can be built in this particular domain. Most of the data, and most of the theoretical views presented were compatible with our view, sometimes clearly supporting it or meeting its predictions.

## References

Arsenijević, Boban. 2008. From spatial cognition to language. *Biolinguistics* 2, 3–23.

Bever, Thomas G. 2009. Remarks on the individual basis for linguistic structures. In Massimo Piattelli-Palmarini, Juan Uriagereka & Pello Salaburu (eds.), *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*, 278–299. Oxford: Oxford University Press.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2, 113–23.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008a. Approaching UG from below. In Uli Sauerland & Hans Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax–Semantics*, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008b. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46, 621–646.

Gärdenfors, Peter. 2000. *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–79.

Hinzen, Wolfram. 2008. Succ + Lex = Language? In Kleanthes K. Grohmann (ed.), *InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces*, 25–47. Oxford: Oxford University Press.

van Lambalgen, Michiel. 2008. Tense, aspect and parallel architecture. Paper presented at the workshop *Biosemantics: Semantics within the Biolinguistic*

*Program*, Leiden and Amsterdam. [Universiteit Leiden and Universiteit van Amsterdam, 12–14 September 2008.]

Tomalin, Marcus. 2007. Reconsidering recursion in syntactic theory. *Lingua* 117, 1784–1800.

Myachykov, Andriy, Michael I. Posner & Russel S. Tomlin 2007. A parallel interface for language and cognition: Theory, method, and experimental evidence. *The Linguistic Review* 24, 457–475.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar: Handbook in Generative Syntax*, 281–337. Dordrecht: Kluwer.

Roeper, Thomas & William Snyder. 2005. Language learnability and the forms of recursion. In Anna Maria Di Sciullo (ed.), *UG and External Systems*, 155–169. Amsterdam: John Benjamins.

*Boban Arsenijević*  
*Universitat Pompeu Fabra*  
*Formal Linguistics Research Group*  
*Carrer Roc Boronat, 138*  
*08018 Barcelona*  
*Spain*  
*boban.arsenijevic@upf.edu*

*Wolfram Hinzen*  
*Durham University*  
*Department of Philosophy*  
*50 Old Elvet*  
*Durham, DH1 3HN*  
*United Kingdom*  
*wolfram.hinzen@durham.ac.uk*

# Analysis by Synthesis: A (Re-)Emerging Program of Research for Language and Vision

Thomas G. Bever  &  David Poeppel

This contribution reviews (some of) the history of analysis by synthesis, an approach to perception and comprehension articulated in the 1950s. Whereas much research has focused on bottom-up, feed-forward, inductive mechanisms, analysis by synthesis as a heuristic model emphasizes a balance of bottom-up and knowledge-driven, top-down, predictive steps in speech perception and language comprehension. This idea aligns well with contemporary Bayesian approaches to perception (in language and other domains), which are illustrated with examples from different aspects of perception and comprehension. Results from psycholinguistics, the cognitive neuroscience of language, and visual object recognition suggest that analysis by synthesis can provide a productive way of structuring biolinguistic research. Current evidence suggests that such a model is theoretically well motivated, biologically sensible, and becomes computationally tractable borrowing from Bayesian formalizations.

## 1.    The Problem

It is a commonplace that perception is in part constructive (e.g., James 1890). The computational mind takes imperfect, blurred, and continuously varying input and reports out discrete representations. The corresponding empirical problem for language exists in several dimensions — phonetic, lexical, phrasal, propositional, and semantic. In each case, the surface input data are insufficient to account for all of what is perceived and used as discrete categories. A large part of the problem derives from the fact that each language is different in its details and there is no computationally tractable upper bound on the number of possible utterances to be perceived. Thus, each level of the perceptual process must involve a creative component, tuned to each input utterance. We review an old solution to this problem, which is gaining new currency because of advances in behavioral, computational and neurobiological research. This solution, 'analysis by synthesis' (AxS), combines hypotheses about the input with the computational re-creation of the input, as a way to combine the contributions of perception and computational reconstruction. We sketch some of the old and new evidence that

enriches this model, and outline a set of research questions that are now becoming salient, in part answerable today, and that set an agenda for future research.

Why should a discussion of this algorithm be of any interest for biolinguistics? The biolinguistic program is rooted in the desire to unify the theoretical foundations of linguistic research with the material infrastructure provided by biology, and especially neurobiology. The goal of this unification is to develop an integrated and explanatory account of how the human brain makes the attributes of the faculty of language possible. This is a laudable goal — but it must be acknowledged that we have very little understanding of how any aspect of speech and language is computed/represented in the nervous system (Poeppel & Embick 2005). There exist interesting correlative insights (of the granularity 'brain area x is typically implicated in function y'), but very little of any serious explanatory depth. It is our contention that an architecture such as AxS provides a way to develop and explore linking hypotheses between the representational architecture of the language system and the psychological/neural mechanisms that form the basis for computing over the hypothesized representations.

A critical feature of the AxS architecture is that it combines statistical pattern recognition, symbolic generative processes and hypothesis confirmation (for example, of the form 'compare the predicted pattern to the actual input, calculate the error, iterate the process until the error is minimized'). These different subroutines that jointly constitute the AxS architecture are gaining support in various areas of language research (Poeppel & Monahan 2010) as well as other areas of perception, notably vision (Hochstein & Ahissar 2002, Yuille & Kersten 2006), and we therefore are optimistic that pursuing AxS (an approach that is broadly consistent with current approaches to Bayesian inference in perception) as a research strategy might be fruitful in studying biolinguistics in a real, practical sense — that is, merging biology and linguistics in the service of one particular problem in perception and comprehension.

## 2. The Re-Birth of Analysis by Synthesis

Consider a simple example:

(1)     Aywannaeate_~dr~nsuPrsftayskriyme~iDay~mz

We hear something like the representation in (1), corresponding to a continuously varying acoustic waveform, but we automatically render it internally as something like the array in (2).

(2)     Phonetic:        Ay w o n a I t t e n dr n s u p er s f t ay s k r I m e n I t ay m s
        Lexical:         I wanna eat tender and super soft ice cream many times
        Phrasal:         [I want [to eat [[[tender and] super soft] [ice cream]] [many times]]
        Propositional: I = agent, want/eat = (double-verb) predicate; more = predicate
                        modifier; ice cream = patient; tender and super soft = ice cream
                        modifier; many times = modifier of predicate
        Semantic:   (yum yum?) …

How does this happen? A great deal of attention has been given to the ostensible initial stage — acoustic mapping onto phones, phonemes, syllables, and words. The emerging theory was (Liberman *et al.* 1967) and for a long time has been (for review, see e.g., Galantucci *et al.* 2006), the 'motor theory of speech perception' (a perspective that continues to receive a lot of attention in the cognitive neuroscience literature, for better or for worse, and where any motor cortex involvement tends to be interpreted, erroneously, as support for this view). On this theory, flowing speech is perceived as intended phonetic-motor articulatory gestures by way of internalized regeneration of the gestures that could have gone into producing the speech. This model called on the AxS-framework outlined earlier by Halle & Stevens (1959, 1963), as a general architecture for *integrating initial analysis of input information with constructed interpretations of it*. Their model aimed to address phenomena that involve a derivational synthesis of the output form from an input, by way of a series of computational steps. For example, the following phonological rules of English must apply in a specific order just to account for the relation between the intended and perceived word /tender/ and its actual phonetic/acoustic form:

(3)   i.     Nasalize vowel before a nasal consonant.
      ii.    Drop a nasal following a nasal and before a homorganic consonant.
      iii.   Lengthen a vowel before a voiced stop consonant.
      iv.    Neutralize voicing in a stop consonant following a stressed vowel and before an unstressed vowel.
      v.     Delete short unstressed vowel to zero before final /r/.
      vi.    Lengthen final /r/ (syllabify it) following a consonant.

This series of rules takes the word /pander/ to [paa~DR] in six easy steps. It is significant that each of the separate rules has broad application in English, not just for the particular word. Thus, it is a consequence of the separate rules that they pile up in a particular order for cases that combine their effect. The crucial importance of such a derivation is brought out by the contrast with the word /panter/ which appears as [pa~DR]. The crucial fact is that the phonemic difference between the two words is the consonant /t/ vs /d/, but the phonetic difference is conveyed only by the length of the first vowel. Recovering the underlying phonemic form from the phonetic form is of course possible by way of a complex set of pattern recognizers — for example, 'if a vowel is nasalized, assume it is followed by a nasal homorganic with the following consonant'. But such surface pattern recognizers become increasingly complex as the derivational processes mount up. In this case, the ultimate pattern input is roughly (in words); if a long vowel precedes a tongue flap before a syllabic element, then assume that the flap indicates a D, otherwise a T. Such 'rules', of course, miss generalizations that characterize the phonology of the language (for example, a 'different' rule would be required to disentangle [luu~BR] from [lu~BR] (/lumber/ vs. /lumper/, and still another rule to distinguish [lii~KR] from [li~KR] (/linger/ vs /linker/).[1]

_____

[1]     But note that phenomena such as these are a serious challenge to non-rule based phonolo-

There are important consequences of computational derivations mediating the relation between an internal representation and a more accessible representation. In particular, they show that *it is computationally intractable to go directly from the more concrete to the more abstract representation by way of filters or other kinds of 'bottom-up' triggering templates.* This feature of language has been understood for more than a century. Thus the levels of representation internal to each component of a sentence are ordered from most abstract to the more superficial. The above example shows this for the phonological → phonetic component. A similar property holds for many models of syntax. In older terms, this is because every sentence has an 'inner' and 'outer' form (cf. Wundt 1900, Bloomfield 1914): Discovering the 'inner' form from the outer form only is computationally prohibitive if feasible at all (see below).

In classical generative grammars, there is an 'underlying' structure, which represents the basic structural relations between constituents and a set of processes that map that structure onto a surface organization of phrases. The puzzle for psychologists and learning theorists has been the great difficulty in relating the two levels by analyzing the outer form and attempting to derive the inner form from it. It is fairly clear why this would be difficult in the case of discovering the underlying forms in the phonological example — and would lose the language-specific generalizations. Similar problems arise in disentangling the inner form of syntactic expressions that appear similar on the surface, for example (4):

(4)     John was eager enough to help.
        John was likely enough to help.
        John was surprised enough to help.
        John was forced enough to help.
        John was strong enough to help.
        John was easy enough to help.
        etc.

In a derivational system, each of these forms has a distinct inner form ascribing different roles to John and different relations between the apparent main predicate and the complement. Chomsky & Miller (1963) noted that the structural result of grammatical processes is that they map a complex hierarchically organized propositional representation of meaning onto a linear sequence. Ostensibly the linear form is unidimensional, although intensive pattern recognition processes may extract several skeletal dimensions, such as 'words', 'phrases', 'intonational units', and so on. But *ultimately some critical information remains unavailable* in the serial signal — in the above examples, the actual syntactic / semantic relation between the apparent subject (John) and the predicate.

Halle & Stevens' conceptual architecture articulates the derivational

---

gies such as Optimality Theory; however, the basic point we are making would hold in the context of an optimality-theoretical analysis, since that analysis would have to be fairly complex to take the facts into account — there is still an abstract computational system mediating the relationship between the lexico-phonological structure of the words and the phonetic output.

processes involved in an AxS model into several logically organized steps.

(5)   A.   Extract a skeleton of the input based on passively recognizable cues.
      B.   Access a derivation that fills in the missing parts of the skeleton.
      C.   Match the output of A to the representation in B.
      D.   If C is successful, confirm the representation from B as the underlying form.

The 'guesses' are generated based on the early skeleton, and trigger the derivation in B. This mapping from template-based guesses underscores the 'hypothesize and test' nature of the AxS algorithm, consistent with the TOTE model that launched the cognitive revolution in the mid 20th century (Miller *et al.* 1960). Halle & Stevens noted that this scheme involves reconstructing the derivation underlying the phonological system, akin to the production of an actual motoric or acoustic representation of the input for matching. However, they emphasized that the actual match can be made internally — matching an abstract computational representation of the input skeleton against a corresponding abstract computational representation of the synthesized match for it. This followed the ideas of Jakobson *et al.* (1952) that phonemes and their distinctive features have an independent computational role in the phonology, while also having regular sensori-motor correlates.

A few years later, from an unexpected direction, Ken Goodman proposed a corresponding AxS model for reading (Goodman 1967). His argument was not as specific or explicit, but argued that printed characters are primarily cues to the 'reconstruction' of the actual text. He argued against the complete bottom-up model of reading, on which readers first translate letters or whole words into their corresponding sound, and then applied their auditory language understanding system to the internal auditory representation of the text. He noted that many errors of reading aloud show that the reader (especially the child) is creating (i.e. predicting) representations ahead of the actual text, which generally correspond to the meaning if not the form. For example a child might 'mis'-read (6a) as (6b), preserving the general meaning and most of the actual text.

(6)   a.   The dog was barking aloud.
      b.   The dog was barking a lot.

## 3.   Enter the Motor Theory of Speech Perception

The idea that speech perception involves reconstructing the production plan is most strongly evident at the acoustic/phonetic level. This idea goes back centuries, at least to von Humboldt (1836), and before that to de Cordemoy (1686). But it received relatively little technical development until the middle of the 20th century, sparked by the failure of filtering theories to explain increasingly sophisticated psycho-acoustic data. In the 1950s and early 1960s it was becoming clear that the acoustic signal required reconstructive analysis at the lowest levels. Ladefoged & Broadbent (1957) used artificial vowel stimuli that correspond to

different shaped vocal tracts which set the reference level for the mid range formant of vowels. The reference level was set by its use in the phrase leading up to a critical stimulus 'please say what this word is…'. They showed that a target word bVt, with the vowel roughly /e/, as in /bɛt/, would be heard as /bæt/ if the introductory phrase utilized a high formant structure and /bɪt/ if it utilized a low formant structure. That is, listeners automatically and unconsciously adjusted their interpretation of the vowel by reference to the formant structure of the vowels in the immediate lead-in — they calculated a midrange expectation and interpreted the target vowel in relation to that. Of course, a moment's thought makes clear that we do this all the time: We have no trouble understanding six-year-old children, adult men and women, despite the radical differences in size and shape of the vocal tracts, with large resulting differences in the actual acoustic structure of their utterances. Furthermore, we do this virtually immediately, starting with the first word we hear someone say. The fact that we have perceptual constancy in light of the considerable variation in the input signal is a remarkable property of the human speech perceptual system, one that highlights the difference between human and automatic speech recognition systems, for which this kind of variability in the signal continues to be a show-stopper.[2]

Facts such as this were compounded by the evidence that we 'hear' sounds that are literally not present in the stimulus. Thus, studies using artificial stimuli (the so called 'pattern playback machine') showed that the percept of a final p, t, and k as in /pɪp/, /pɪt/, and /pɪk/ depends entirely on the vowel transition up to the final consonant — indeed, the consonant can be totally lacking, or represented just by a neutral burst of aspiration, and the differentiation is clear: In other words, listeners 'hear' a consonant that in fact is not present — rather it is the vocal gesture leading up to the silence that conveys the shape of the vocal tract as the vowel stops.

Such considerations supported some of the general assumptions underlying the motor theory of speech perception — the view that at the outset, listeners are reconstructing the articulatory gestures of the speaker, and using those as the trigger for the perception of the underlying intended sequence of phones as though they actually occurred acoustically. This theory persists today. Of course, it can always be recast as a pure perceptual 'bottom-up' theory, if one assumes an arbitrarily large number of such filters. In the end, as often is the case, the argument in favor of such a constructive theory is not logically apodictic, it is empirically indicated. Recent attempts to provide a dynamic alternative to a constructive theory involve Bayesian models, in which the initial input is organized into recognized units using the probabilistic extent to which the input represents the units. This kind of model has achieved some success in computer vision (e.g. Fei-Fei & Perona 2005) and in lexical identification in speech (e.g., Norris & McQueen 2008). With the initial goal of recognizing a finite

---

[2]    Recent research (e.g., Lotto & Holt 2006) has shown that the Ladefoged & Broadbent-effect can be achieved simply by preceding the target /bet/ with a high or low filtered noise. This shows that setting the expected mid-range does not depend on actual speech; but it still requires that the listener is using the information to set expectations about the vocal tract of the speaker.

number of objects (e.g., 30,000 visual types; roughly the same number of words), the models achieve some success, within the domain of computational modeling (say, 90% correct). But the problem for whole sentence recognition is different both because of the complexity of syntactic organization even for simple sentences, and because of the indeterminate upper bound on sentence length. Furthermore, unlike constructive models based on grammatical structures, the statistical models generally fail to represent a great deal of what we know to be true about sentences, for example, remote structural properties, structural details of phrasing etc. We return to this below in the discussion of syntactic parsing.

## 4.    Neisser's (1967) Elaboration of Analysis by Synthesis

Halle & Stevens' papers were stimulating intellectually but had little immediate impact on the study of language comprehension at levels more abstract than speech processing. In the speech recognition literature, too, attention turned to the utility of statistical processing models of the Hidden Markov type, where little emphasis was placed on the value of the knowledge of language, whether it is phonological, lexical, or syntactic, or semantic. A notable exception is the remarkable book by Ulric Neisser, *Cognitive Psychology* (Neisser 1967). Neisser reviewed the available evidence showing 'top-down' processing in vision as well as language and other areas of cognition. At the time, the book caused a stir because it was the first programmatic statement that consolidated much of the re-volution against the prior dominant behaviorist views on which perception was primarily a 'filtering' process, from external input to internal representation. As Neisser put it, redolent of William James, "The central assertion is that seeing, hearing, and remembering are all acts of reconstruction, which may make more or less use of stimulus information" (p. 62). But, while given some attention, it did not spark intensive development of the AxS model, and Neisser himself turned to more ecological and contextual concerns as the logical extension of an approach that emphasized constructive influences in cognition.

## 5.    Analysis by Synthesis as a Solution to the Syntactic Generation Problem — Perceptual Strategies

Meanwhile, within the psycholinguistic world, evidence was being developed that generative rules play a role in language not just in phonology but at the syntactic level as well. By the late 1960s, George Miller and students had amassed evidence suggesting that the underlying structures of sentences were computed as part of sentence memory, recognition and understanding. For a time it appeared that the syntactic rules and ordered derivations that they defined could be taken as corresponding to psychological operations. The one-rule/one-operation hypothesis was testable in general by assuming that sentences with more rules involved in their derivation would be correspondingly more complex: A passive sentence should be harder than an active, a passive-negative sentence harder still, and so on. At first this 'derivational theory of complexity' (DTC)

appeared to be supported: But eventual careful study showed that it was not systematically the case (Fodor & Garrett 1966, Bever 1970). Recent research in cognitive neuroscience of language has reopened the debate on the DTC (see, e.g., Marantz 2005). Methodological progress and theoretical shifts suggest that something like a mapping from representational complexity to number of computational steps may be on the right track, and such a perspective is implicitly at the basis of much work in experimental language research. For example, experimental research on lexical structure (morphology) as well as on lexical semantics suggests that structural complexity is associated with changes in processing cost as reflected in both behavioral and neurophysiological indices (see, e.g., Gennari & Poeppel 2003 regarding lexical semantics, where more hypothesized structure correlates with longer reaction times or Fiorentino & Poeppel 2007 and Zweig & Pylkkänen 2009 regarding lexical structure, where neural data from MEG distinguish between simplex and complex words). However, it remains to be shown, either behaviorally or neurologically, that something like DTC is correct at the level of derivational syntax or compositional semantics. Bever (1970) suggested that in an AxS-framework, each syntactic rule can correspond to a mental operation: But the small processing difference from different number of transformations is obscured by the initial input strategies that give a preliminary analysis of the sentence meaning. Thus, the DTC could be true computationally, but not show up in some actual behavioral complexity differences. Bever argued that an initial set of 'perceptual strategies' is necessary in order to establish the equivalent of the input skeleton assumed for the phonological analysis by synthesis scheme. For example, in English almost every finite clause has the surface form (8a), excluding interjections and adjuncts, which corresponds thematically to (8b):

(8)    a.    NP/agr Predicate/agr XP
       b.    Agent predicate other (patient, complement, etc.)

Accordingly, a first pass through most clauses can rely on a scheme that looks for structures like (8a) and maps them directly onto thematic relations like (8b). At that point, the grammar can apply (or have applied in parallel) the set of transformations to 'check' that the initial analysis is consistent with a corresponding derivation and is correct. In many cases it will be, but in selected cases, such as passives, object-clefts or object relatives, it is violated: And it is just those cases that the succeeding 30 years of research have shown to be particularly complex in normal processing, difficult for aphasics and so on. In this regard it is important to remember that the initial semantic mapping of a phrasal sequence onto a set of thematic roles is not itself a syntactic derivation: Thus, even the simplest sentences still requires a constructive component of some kind.

This brings us to a critical question underlying debates about sentence comprehension in general: *Are grammatical derivations computed as part of the processes of comprehension?* The question can be addressed in several aspects, and it is useful for us to clarify our position on them. First, are syntactic derivations correct descriptions of what speakers know when they know a language? This question can be answered negatively, as in remarks by many connectionist

theorists or more recent statistical modelers (e.g., Lappin & Shieber 2007): On these views, grammatical 'rules' and 'derivations' are themselves statistical generalizations over actual instances of utterances — accordingly, an adequate statistical model will actually capture the essence of language structure correctly. At the moment this assertion continues to be a promissory note (startingly like that of Zellig Harris; see papers in Harris 1970). Computational modeling struggles to achieve a modicum of success in assigning correct lexical categories after supervised training (the best claims going from about 85% to 90% in the last 25 years; see Charniak 1997 and Titov & Henderson 2007); less has been achieved in assigning correct tree structures.

Of course, such 'failures' don't look that bad in numerical terms when stacked up against actual linguistic analyses: They are generally incomplete, because they are motivated by circumscribed theoretical issues, not attempts to master the whole grammar of a language at one time. The enduring problem is that there are many systematic facts about sentences that are captured by grammars with derivations, which are not even in the goal set of statistical modeling. A sample example is the full range of phenomena described under C-command constraints (constraints that relate processes in a phrase level to its descendants in a tree): Many grammars that appear to differ greatly, share the corresponding properties (e.g., generative grammar, lexical functional grammar, categorial grammar). It is for reasons like this that *our discussions here presuppose that some form of structural grammar is correct* for the language and for the speakers of the language.

The second question is whether derivations are actually applied during comprehension. This is an empirical question of a different kind: It has preoccupied a small band of psycholinguists for 50 years, since the original work on the 'psychological reality of grammar' started by George Miller and his colleagues. Of course, the most massive data in favor of the role of derivations is the immediate recognition of whether a sentence is grammatical or not, as part of understanding it. These data vastly outweigh any set of experiments. But in addition, we accept the considerable evidence that syntactic derivations are assigned as part of comprehension processes; perhaps not always the most important part in some contexts; perhaps circumvented by memorized idioms in some cases; but we assume that the comprehension system is always prepared to assign a derivation (see Townsend & Bever 2001, Crain *et al.* 2008, and Wagers & Phillips 2009 for examples of some relevant empirical findings).

Finally, we note that several models have grafted Bayesian or other statistical modeling onto an existing grammar. For example, Morgan *et al.* (2010) propose a statistical interpretation of the set of categorial grammatical rules that generate the benchmark trees in the Penn Tree Bank; Riezler *et al.* (2002) make a similar proposal for interpreting Lexical Functional Grammatical Rules. In each case, the linguistic grammar is presupposed, as well as the kind of derivations that it assigns to individual cases. The role of the statistical metric on each rule is to yield sentence structures that approximate their distribution in some corpus.

Given the presence of derivations as part of sentence comprehension, the AxS model meets an obvious puzzle especially at the syntactic level: Sentences stream serially in time word by word, but derivations are computationally

'vertical', with at least entire clauses as their domain. That was true of early syntactic models as in Chomsky's *Syntactic Structures* (Chomsky 1957) or *Aspects of the Theory of Syntax* (Chomsky 1965). But the many recent models actually build sentences up from the most to the least embedded portions, which in English means from the right to the left (Chomsky 1995).[3] This sets what we think of as the logical problem of sentence comprehension: It is serial, but vertical at the same time. Townsend & Bever (2001) address this question directly and argue that it is a further argument for analysis by synthesis. But it also emphasizes that the initial pass must usually have enough information in it to engage at least a preliminary meaning: As they put it, "we understand everything twice", once based on the initial perceptual strategies such as (5A–B) and then again via the actual derivation. They suggest that we do not notice the multiple phases because the second follows the first within a 200 millisecond window, resulting in a representational merging of the two meaning representations. Bever (1992) and Townsend & Bever (2001) also note a general implication of this kind of dual processing: It unifies inductive based comprehension with deductive computation based comprehension. That is, it unifies, or at least binds together, the two main insights of centuries of cognitive science:

(9)    i.    Much of what we do is based on habits accumulated via induction over experiences.
       ii.   But some of what we do is based on novel computation.

### 5.1.    Analysis by Synthesis in Automatic Speech Recognition Systems

Aside from the motor theory of speech perception, early stages of automatic speech recognition utilized AxS procedures. The literature on this is vast, in part because of the practical importance of automatic speech recognition systems. We touch only on an early and current stage of thinking about the value of AxS in speech recognition. For example, Bell *et al.* (1961) applied the method to reduce the search space of phonetic sources of speech spectra. In the succeeding five decades, the field of automatic speech recognition has witnessed the development of many sophisticated filtering procedures, that operate in a 'noisy channel model', Bayesian statistical filters, and so on. Thus, the array of apparent 'direct perception' devices and models has expanded greatly, and converge onto a high degree of success (Huang *et al*. 2001). However, if one looks closely at how these models often work, one sees a 'frozen' instantiation of an AxS scheme (Jurafsky, p.c.). For example the noisy channel model of word recognition includes a generative model of the words-to-waveforms process: When a waveform comes in for recognition, the model checks every possible word string, runs it through the words-to-waveforms process and picks the one that is the closest fit.

The salient difference between this kind of AxS model and Halle & Stevens (1962) is that the model is parameterized at a different level of granularity; instead of modeling the articulatory system, the process keeps a Gaussian model

---

[3]    But see Colin Phillips' work for left-to-right computation, incorporating knowledge-driven predictions to generate potential structure (e.g. Phillips 2003.)

that directly stores vectors representing mean and variance of spectral slices. Furthermore, some models that include a more explicit AxS component continue to be argued as superior to those that do not (Bawab *et al.* 2008). For our purposes, the important conclusion is that despite enormous computing power of today's machines and the development of powerful statistical tools, an AxS component for speech recognition continues to be critical, if typically implicit and unstated in descriptions of the systems.

## 6.    New Data Bearing on AxS

### 6.1.    *Audiovisual Speech Perception*

Unexpected recent support for the AxS approach to perception derives from data on the multi-sensory processing of speech. Until recently, speech perception was primarily studied from a purely auditory perspective, and, obviously, any successful theory of speech perception must account for the range of phenomena based on processing of the acoustic signal alone, since listeners perform well in absence of any additional cues (e.g., listener is turned away, has eyes closed, is in the dark, is blind, is listening over the phone, the message is on a totally unfamiliar topic, etc.). That being said, a significant proportion of our communicative interactions occur face-to-face, and it has become a topic of considerable interest to evaluate how the senses interact and/or 'merge' during perception. The standard view is that facial cues provide additional information that reduces uncertainty (in an information-theoretic sense) and augments the perceptual interpretation suggested by the audio signal. On such a view, an audio signal (say, a syllable) activates possible targets (e.g., as in the cohort model (Marslen-Wilson & Tyler 1980 or the TRACE model of McClelland & Elman 1986) and the associated video signal (say, the face articulating the syllable) provides convergent input, pushing the activated nodes closer to firing threshold. The senses yield independent but convergent data from the input and the processing streams are merged to elicit the suggested perceptual analysis.

Some new experimentation suggests an alternative (or additional) perspective on AV speech. Van Wassenhove *et al.* (2005) presented listeners with AV syllables — including both audiovisually congruent and conflicting information as in McGurk & MacDonald (1976) — and recorded the ERP while viewers/listeners reported what they perceived. The major evoked responses elicited by auditory stimuli, the N1 and P2, were modulated by the presence of the facial information in surprising ways: The timing of these responses (e.g., the peak latency) changed as a function of how informative the facial cues were — in a facilitatory direction. Highly informative facial information (and hence articulator information) led to significantly shorter response latencies. Because in these utterances the movement of the face always preceded the audio signal (as is typical of natural utterances, that is, the articulators have to move prior to sound emerging), it was argued that the facial information predicts possible audio signals. Since they varied the facial information parametrically, they were able to show that there appears to be a systematic relation between the information that

the face predicts and the temporal savings. The best explanation was argued to be an AxS approach, in which the visual signal elicits 'guesses' (akin to the templates mentioned above) for possible sound targets; these hypothesized targets are then synthesized in a derivational step and compared to the actual input; close matches yield strong facilitation.

The response profile reported by van Wassenhove *et al.* (2005) was recently replicated and extended by Stekelenburg & Vroomen (2007) as well as by Arnal *et al.* (2009). The former showed a similar response facilitation for AV speech, but were able to show that such a facilitation can also be observed for other causally, predictively related audiovisual events. For example, the movement of a hammer towards a surface predicts a sound of a certain type in a specific temporal interval; interestingly, the neurophysiological response to the sound alone is significantly longer than to the audiovisual event. This suggests that the predictive relations of this type are not speech-specific, and that an AxS approach might be extended to perception more generally (reminiscent of the systematic arguments made by Neisser 1970). The input signal from one modality suffices to trigger 'guesses' (perceptual hypotheses, the induction part of AxS) that make contact with the abstract internal representations that permit derivation of the possible targets (the synthesis part of AxS). A critical issue is, naturally, what the format of representation is that mediates between the initial guess and the derivation/synthesis of the target. For speech, there exist well motivated representational theories that can be used, say the notion of distinctive features. It is less clear today how non-speech information is encoded and represented.

### 6.2.    *Cognitive Neuroscience Data on the Perception-Production Link*

Recent data from various corners of the cognitive neurosciences have reignited interest in the idea that there is a tight mapping between perception and action. Although the motor theory of speech perception has played a dominant role in theorizing on that topic, the majority of experimental approaches to perception focused on feed-forward approaches, by and large sidestepping the issue of a link between perception and production. However, neurobiological data deriving from the arsenal of contemporary approaches have supported, at least to some extent, the view that brain areas typically associated with the generation of output play some role in the analysis of the input. These new data raise the question of whether activation of motor (output) areas merely reflects associative mechanisms that link perceptual and motor areas (for example, watching track & field is — unsurprisingly — related to knowledge of how legs work in running, say), or whether the motor activations play a genuine role in the analysis of the input. If these output related activations provide a real (necessary) contribution to perceptual analysis, a further question is whether analysis by synthesis is the type of algorithm that is instantiated by this pattern of activation.

Importantly, it is not established whether motor activations play a causal role in perceptual analysis, all claims to the contrary notwithstanding. On the positive side, both hemodynamic imaging and electrophysiological recording have demonstrated robust contribution of motor cortical activations in various perceptual tasks. For example, Wilson *et al.* (2004), using fMRI, have provided

data showing motor cortical activation during passive speech perception. Similarly, Skipper *et al.* (2007), also using fMRI, document the activation of motor areas during the viewing of audiovisual speech. Both sets of results have been interpreted to support the view that these areas contribute to speech comprehension. Using electrophysiological techniques, such as EEG and MEG, other investigators (e.g., Pulvermüller *et al.* 2006; see review by Pulvermüller & Fadiga 2010) have shown that electrophysiological responses localized to motor areas are active remarkably early in the processing stream (say within 200 ms), once again suggesting that the neuronal tissue associated with the generation of output is active during the time interval typically associated with perceptual analysis. D'Ausilio *et al.* (2009) report selective interference of the discrimination of CV syllables when the corresponding motor areas are temporarily inactivated by Transcranial Magnetic Stimulation, a technique that generates temporary localized lesions. Cumulatively, the electrophysiological and the hemodynamic imaging data provide positive evidence for the conjecture that motor areas are *somehow* involved in perception. However, interpreting such activations as evidence for an AxS view is rather more complex: It would require that the hypothesized perceptual targets are internally synthesized, that is, that there is a deductive, derivational computation that precedes the comparison of the input signal to the internally generated candidate representations. The data that are available to date have not been analyzed in the context of such a perspective.

It is also important to bear in mind that there are data which provide a challenge to the simplest possible story outlined here: The findings from brain injuries, by and large strokes, do not support the hypothesis that motor areas in the frontal lobe are required for successful perception; at least this is true for the case of speech perception (for review, see Hickok & Poeppel 2007), and it is unclear to what extent motor areas are critical for action perception in other domains. The simple story one might it envision is like this: Motor areas generate action plans and ultimately instantiate the action by triggering the motor neurons that drive the musculature. These frontal areas are connected to the posterior perceptual cortical fields, and their direct anatomical connection suggests physiological co-activation (via efference copy). On that view, the frontal areas can provide the substrate to generate guesses about the output that are then fed back to the posterior areas that evaluate the input. This, for example, would be a reasonable interpretation of the DIVA model for speech production (Guenther 2006). However, the lesion data make such a straightforward interpretation very problematic. It is simply not the case that lesions to motor areas lead to catastrophic consequences (or any consequences) for perceptual analysis. Data from transcranial magnetic stimulation also provides mixed results on the involvement of frontal motor areas at the lexical level. Research by Rumiati and her colleagues — for example Papeo *et al.* (2009) — documents that the processing of verbs denoting motor actions is not disrupted by stimulating the corresponding motor areas during comprehension. In sum, either this means that frontal areas play no critical causal role in perception, or that, in fact, there exist posterior cortical areas that are involved in the programming of production. This latter perspective is the one endorsed for the processing of speech by Hickok & Poeppel (2007), where it is argued that a cortical field at the interface of the

temporal and frontal lobes provides the critical substrate for mapping from input representations to output representations.

In the present context it is important to mention one frequently raised putative mechanism to link perception and action. There exists a class of neurons that has, in the recent literature, attracted considerable attention and been invoked as the cellular substrate from phenomena ranging from the evolution of language to empathy to theory of mind. These so-called mirror neurons (Rizzolatti 2005), active during the execution of an intentional action as well as the observation of that action, have been argued both in the professional and popular press to form the neural substrate for the 'understanding of action'. Cells with these particular characteristics are observed by many labs, and the nature of the data is not disputed. On the other hand, the interpretation of what these cells do is entirely unclear. A recent review of the mirror neuron literature (Hickok 2009) suggests that, even for nonhuman primates, the interpretation that mirror neurons constitute the basis for the 'understanding of action' is much too optimistic. And, worse, simply unsupported…

That being said, one could imagine a narrow and computationally specific role for mirror neurons, especially those documented for auditory cognition (Kohler *et al.* 2002). In particular, if there exist cells in the frontal, parietal, and temporal cortices that fire during the mouth movements, and if these same cells fire during the observation of the same type of articulator movements, one could imagine that such cells play a role in mediating the 'currency' that the brain has to use in translating back and forth between generating speech output and analyzing speech input. If, say, the currency of speech sound processing is the 'distinctive feature', then cells that facilitate the mapping of such computational primitives both to the output side (articulator configuration) and to the input side (acoustic template of a feature; cf. Stevens 2002) would be extremely useful. The utility of such cells notwithstanding, their existence would obviously not suffice as an argument that AxS is an architecture that organizes the processing. In short, mirror neurons could, perhaps, be adopted and adapted to play an important role in how analysis-by-synthesis is instantiated; however, it will be important to find a circumscribed, narrow, computationally explicit role. Invoking these cells to solve everything from evolution to impotence is not helpful, even if amusing. The main message of this section, even if a bit messy, is this: there is convincing evidence that motor cortical areas are activated during perceptual tasks. And in some of the cases, the so-called mirror neurons are implicated. However, it is not clear that we are in a position to argue that these particular output-related cells form the basis for the analysis-by-synthesis approach. That offers one elegant and simple solution, but the data do not compel one to this view alone.


## 7.    Analysis-by-Synthesis in Visual Perception

Interestingly, research on visual object recognition has, in the last few years, made contact with the concept of AxS as well. As discussed above, the AxS concept was first articulated in the context of speech perception, by Halle & Stevens. It was subsequently elaborated by Neisser, and connected in important

ways to the formulation of the motor theory of speech perception of Liberman *et al.* But, curiously, the concept has played no major role in any aspect of perception, save certain parts of psycholinguistics, for a long time.

Research on computational vision, and in particular on visual object recognition, has '(re)discovered' a form of AxS because three closely related concepts have played a prominent role in recent work, concepts that in turn form the basis for AxS. One stream of research that has been productive and very informed by data from systems neuroscience and single unit recording is the notion of *predictive coding*. It is now well established that there is a robust predictive aspect to visual perception; the visual system 'expects to see' specific shapes or other visual attributes (motion, color, texture, etc.) and predicts properties of the *anticipated* visual targets. Predictive coding is observable in the neuronal firing properties of neurons in various visual cortical fields.

The second strand of research that has been influential in computational vision is Bayesian perception. The Bayesian conceptual infrastructure links notions of conditional probability, the ongoing perceptual data, and the priors. Calculating the posterior probabilities involves a prediction of the anticipated image; calculating the prediction is closely related to the notion of a derivation of a candidate target. Research on 'vision as Bayesian inference' makes explicit use of the analysis-by-synthesis architecture (Yuille & Kersten 2006).

A third area of research has focused on the calculation of the prediction error, and how to use that error in improving the next processing step and updating the current representation. This work has been able to develop detailed neurocomputational models that show how the error is used, in studies ranging from arm movement control to reward control. Importantly, brain imaging data and electrophysiological data have been used successfully to support the hypothesis of predictive coding, Bayesian analysis, in visual object recognition. These data from a different domain of inquiry are important to linguistic research because they point to *generic computational mechanisms that neural systems can exploit in the service of recognition tasks*. If models from vision — perhaps even tested neurophysiologically in animal models — provide data for the subroutines of AxS, we stand to learn something about the implementation of such an algorithm for language comprehension as well. Minimally it suggests that the 'parts list' to build such an algorithm exists.

One example of how the AxS idea might work in visual object recognition is provided by the work of Moshe Bar (Bar *et al.* 2006, Bar 2007, Kveraga *et al.* 2007). Bar *et al.* build on the fact that visual scenes are broken down into different spatial frequencies in the periphery and the afferent visual pathway. One part of the pathway, a 'channel' that happens to be particularly fast in terms of its analysis and transmission speed (the so-called magnocellular channel), is specialized for low spatial frequency information, basically conveying a coarse image of the shape of an object, based largely on contrast information. The high spatial frequency, detailed information is carried by an anatomically separate, slower channel that projects to different areas of the visual cortex (inferotemporal cortex). Now, Bar *et al.* hypothesize that, confronted with a retinal image, the fast 'coarse' channel projects to frontal areas and triggers predictions based on the coarse shape information (cf. in language, the initial templatic guesses). These

guesses are then elaborated (synthesis step) and compared to the more detailed, spatially fine-grained information that arrives in the temporal lobe somewhat later (parvocellular projections). Crucially, this model requires information processing channels whose processing is offset in time — and, conveniently, there is good evidence for such differences in processing times in the visual system. Interestingly, there is some evidence that auditory processing also proceeds on different time scales (see, e.g., Poeppel 2003 for discussion), suggesting that the neuronal infrastructure for a similar scheme might exist for auditory cognition especially relevant for binding different levels of linguistic representation.

## 7.1.  *A x S in Vision and Language: Two Choices*

The reader may have noticed that the three background factors of computer vision might argue for Bayesian models rather than the AxS architecture. Predictive coding, Bayesian modeling, and error-based-correction correspond to the three main components of the AxS architecture: Statistically justified initial hypotheses (aka 'perceptual strategies') can (and probably now should) be modeled using Bayesian approaches to measure the probability of a particular pattern fitting the input; at the same time, as it applies serially, the pattern probability makes several kinds of predictions, namely the structure that will appear on the surface, and how the entire sequence is mapped onto a semantic representation; the role of the 'synthesis' component is to compute a derivation that fills out the analysis, and provides a surface string that checks for surface identity. When there is an error in that, a different lesser hypothesis is chosen as the input pattern, with a repeat of the corresponding derivational check. The application of the predictive component makes it possible to engage the process near the beginning of each major syntactic unit (e.g., a clause) without having to wait for the serial input. This enhances the predictive aspect of the model, indeed it gives the combined role of initial pattern and derivation assignment a strong basis that can turn much of the comprehension of a sentence into a confirmation rather than perceptual analysis process.

Another parallel between a Bayesian framework as developed in computer vision and AxS for language is the role of 'generation' of complete representations. As noted above, the task in computer vision is taken to be to organize input fractional representations into organized arrays that correspond to some interpretable visual form. Various attempts at making this process efficient involve positing hierarchically layered organizations, each successively more precise. In that sense the Bayesian statistical generator provides a notion of 'derivation' in matching each input array to its best fit object.

Thus, we see no incompatibility between the AxS architecture and the role of Bayesian modeling. The difference in the case of language is that, unlike vision, there is a great deal known about what each level of representation is made of and how it is related to its hierarchically adjacent levels. Phonemes are parts of syllables which are parts of words which are parts of phrases which are parts of clauses which are parts of sentences… Thus, the notion of 'generation' of a derivation that links these different levels for each sentence is typically more constrained in linguistic than visual models. Most important, as we noted, such

generative models also incorporate processes that may explain a range of linguistic phenomena other than mere representation of each string.

In the end, our view of this aspect of the current situation in computer vision is that an architecture like AxS may eventually lead to better motivated specification of what visual features are directly relevant for vision and how they are hierarchically organized, ultimately leading to a situation like that in classical and today's psycholinguistics. The model can be taken as framing predictions about relations between scenes, ease of perceiving a given scene, ease of visually grasping how one scene blends into another, etc. This possible effect of the success of the AxS model in language will be a most satisfying result.

## 8.    Today's Research Questions

As the general model is taken increasingly seriously, AxS raises many theoretical and empirical questions that have only scantily been addressed up to now. Here are a few that may serve as guidelines for some next steps in research on the model and the problems it seeks to solve.

### 8.1.    Is the 'Motor' Activation Abstract or Concrete?

Halle & Stevens proposed that the synthetic component that regenerates the derivation of the input, results in an 'abstract' motor code, not the actual motor actions. In the case of phonology, this might be best thought of as a series of sets of linked distinctive features that represent the phonemic description without specifying detailed acoustic or motor correlates. The 'motor theory' in principle suggests a more actuated motor program, but it could still be viewed as an 'abstract' but neurologically organized motor program for articulation, not actuated in real articulatory movements (as in some of Liberman's writings). Some of the questions are a bit hypothetical given today's methodological limitations: Thus, the 'motor program' could consist of the activation of a string of phonemes in the motor cortex that go nowhere, or that go as far as the basal ganglia but no further, that are sent as an efferent copy to the auditory cortex, and so on. Of course, the notion in the motor theory of 'reconstructing' vocal gestures implies at least an internal representation of actual vocal movement, but one could envision that the gestures themselves are actually represented as internal programs. All this relates to the next question, namely:

### 8.2.    Is the Resynthesis of a Derivation Related to the Recomputation of the Linguistic Derivation Only — Or Does It also Include Activation of the Extralinguistic 'Action' Indicated by the Sentence?

Some recent research suggests that specific linguistic representations in the motor area of the cortex are activated shortly after the corresponding perceptual areas are activated. This has been shown for certain kinds of lexical access (Canolty *et al.* 2006, Pulvermüller *et al.* 2006, and Skipper *et al.* 2007). While the behavioral measure (e.g., lexical decision) may itself stimulate motor activity the results are

initially consistent with the AxS model (see section 6.2 above). The sequence of activation from perceptual to motor areas could correspond to the computation of the initial perceptual representation followed by the 'checking' motor representation. A more radical view in a substantial body of today's literature focuses on evidence that the motor activation that plays a role in comprehension, is actually activation of the actions that the meaning of the sentence indicates. Stroop phenomena are an old demonstration of the interaction of a decision or action in the face of conflicting signals: Given an instruction to choose and name a word in capital letters, the choice between /SMALL/ and /big/ is harder than between /small/ and /BIG/: The effect of congruence of the choice and the percept suggests to some that the percept itself activates the action which then can conflict with activating the correct choice. If motor programs for actual actions are activated during comprehension, this makes the next question about semantic interpretation a critical one:

### 8.3.   How does AxS Work at the Level of Meaning?

If the syntactic system reports out a semantically organized meaning, that still needs to be interpreted in term of actions, the 'motor' output would be an interpretation into (possibly an abstract representation) of the action to be taken. Consider a simple example: (10a) is specifically a request for information about the hearer's knowledge of the room's window-opening potential. But it would ordinarily be mapped onto a world in which the reason to request such information is actually interpretable as a request to do something about opening the windows, or at least changing the air quality in the room somehow. So, the utterance has to be interpreted in light of why the speaker might have generated it, that is, it is re-synthesized from its context via a combination of social knowledge, cultural norms, and so on.

(10)   a.      It's stuffy in here. Do these windows open?

Acceptable responses are outlined in (10b).

       b.      "Unfortunately, no."
               i.  [hearer opens a window, breaks it with a hammer, etc.]
               ii. [hearer turns down the thermostat, turns on a fan, etc.]

       In other words, the hearer has to have generated the underlying source of the meaning of the speaker's question in order to respond to it properly. There is a body of research on such indirect requests, mostly carried out via psychological experimentation. The usual question is whether special computations are needed to extract the indirect request from a literally interpreted sentence form or whether there are 'direct' interpretive mechanisms: The literature is divided on this. However, the problem with most of this research is that it uses conventionalized forms for indirect requests, such as in (10c).

       c.      Can you open the window? Do you know the time?

Can you tell me how to find the railroad station?

Since everyone agrees that such forms are structurally set, it is no surprise that in some cases they do not involve extra processing. Research on unconventional indirect requests, such as (10a) is required to learn how pragmatic inferences are computed, (and whether there is computational or neurological evidence for an AxS component in their computation). Recently, Boulanger *et al.* (2008) report some evidence bearing on this: For example they found activation in corresponding motor areas when subjects perceived metaphorical sentences, such as 'John grasped the idea'. However, this still may only show concurrent activation of the lexically coded motor areas, not necessarily directly implicated in comprehending the metaphor.

### 8.4.   How Is the Initial Linguistic Input Categorized so Quickly in Ways that Lead to Correct Derivations Almost All the Time?

This is the equivalent of rapid error detection in the corresponding stage of vision models we discussed. This mystery exists at every level of linguistic representation. Surprisingly, it may be easiest to understand and explain this at the level of syntax: How is it that the initial structural analysis can simultaneously have two critical immediate results?

(11)   i.      Create a surface-to-semantic representation that is (at least close to) correct.
       ii.     Trigger a derivation that is correctly directed to generate the input surface form.

The fact that the initial semantic representation is almost always correct (enough) follows from (or is causally related to) several facts that seem to be universal across languages (see above).

(12)          Every language has a Canonical Syntactic Form (CSF).
       i.      The CSF is the most frequent surface form (e.g., in English, 'NPx Vx [XP]'; in German, '[XP] Vx…'; in Turkish, '<NPx> V <XP>' (<> indicates free word order); in Japanese, 'NPx [XP] V').
       ii.     The CSF has an overwhelmingly dominant mapping onto semantic relations (e.g., in English, 'NPx = agent/experiencer, Vx = predicate/state…').
       iii.    The cases of a surface CSF in which (ii) is not true can nonetheless be initially understood via a misparse based on a simpler form (e.g., passives can be initially misunderstood as complex predicate constructions: 'Athens was attacked by Sparta' can be initially parsed as 'Athens BE (Pred = 'in the state of being attacked by Sparta')').

Clearly, languages can have a few exceptions to the Canonical Form. In English, the main exception is *wh*-fronting as in object-first clefts, interrogatives, and object relatives: Generally, such constructions are signaled by unique mor-

phemes (/who/) or a unique sequence ('NP, NPx Vx'). In general, it is arguable that attested languages are those computationally possible languages that are filtered by the requirement of a CSF (see, e.g., Bever 1970, 2009, and section 8.5 right below for discussions of the role of acquisition in this filtering process).

The second feature of the AxS process at the syntactic level is the presumed accuracy of triggering a correct derivational process to provide a complete syntactic description. In the cases of a full CSF, the correct derivation is close to the initial parse, so there is relatively little mystery. The deeper question arises in explaining how a non-conforming CSF nonetheless receives a correct derivation fairly rapidly. The first part of the answer is that in fact there is a noticeable delay in arriving at the correct derivation — thus, passives in English are fully compre-hended more slowly than actives. The second part is that the initial felicitous misparse in such cases, provides a schema that renders the correct thematic relations, despite the syntactic misparse.[4] It is often thought that verb final languages must falsify the idea that an initial stage of comprehension can proceed based on canonical patterns — if the verb has not been presented, how can arguments be processed in relation to each other? Prima facie considerations like this could be taken as even more evidence for AxS. However, the initial input patterns can include as yet unfilled variables: For example, in Japanese, when a noun with –*wa* is encountered, it triggers the analysis of the noun as a subject/ agent, in relation to an object noun that has already preceded it or that follows it. For English speakers, it may seem odd to posit an a thematic role for a noun phrase before the verb is present. But in fact, English speakers do this easily, as for *John* in:

(12)   John seemed to be upset by Bill.

This example is significant because — in theory — it involves successive assignment of first agent role and then experiencer and then patient role and then experiencer again to John, all before or just as the verb *upset* is encountered. That is, the 'synthetic component' of the AxS scheme must closely follow the analytic pattern templates serially, with as yet unspecified or changeable variables as part of the derivational computation.

## 8.5.    *What Is the Role of AxS as a Model of Learning?*

What is the role of AxS as a model of learning? We have emphasized the perceptual problems that AxS seeks to solve — the inadequacy of surface input to quickly determine the entire inner structure of a sentence or object. This is a problem for adults who have already mastered knowledge of their language and visual world. Now consider the problem of how the child learns or discovers the inner representations of her language and physical world. This is an even greater mystery, especially in light of how quickly the child learns from relatively impo-verished input. The common solution is that the child's search space is critically

---

[4]    See Townsend & Bever (2001), who detail how the series of operations that take the correct thematic relations as input can derive the correct surface form.

reduced by innate expectations and parameterization of what is to be learned: On this model it only takes a small amount of data to resonate with a particular innate structure, or to 'set' a particular parameter — learning consists essentially of throwing a bank of pre-wired switches to conform to the shape of the input.

Recently, Bever (2009) has argued that this scheme is, at best, an abstract description of the boundary conditions on the minimal data that the child must be exposed to for learning about its language and world. The description says nothing about the actual mental activities that the child is carrying out in the process of learning to use its language. Bever elaborates on some initial ideas in Townsend & Bever (2001), that the AxS model may be reconfigured as a model of acquisition: On this model, the child builds up statistical generalizations about the structure of his language — for example, in English that all sentences are of the basic form 'NP V(agreeing with NP) (XP)', where the first NP is the agent of the predicate. The child then accesses its innate grammar-building processes and structures (e.g., phrase structure creation) to provide a derivation for the generalization. This is critically triggered by experiencing the fact that certain sentences that seem to conform to the semantic generalization actually do not (as in passive sentences, raising sentences, and so on).

In this case what is 'synthesized' is a kind grammatical derivation itself, what is 'analyzed' is the surface form and its regular semantic interpretation. This model is an instance of a traditional model of learning and problem solving — an ongoing cycle of inductive hypothesis formation and deductive testing of it. Indeed it is redolent of Miller *et al.*'s (1960) TOTE model of learning. Bever draws a number of factual conclusions that should be true if this model is correct. For example, the model requires that languages present salient generalizations of sufficient regularity to build up patterns from sparse input. This is true of all attested languages, a fact often noted but not attended in relation to its implications — that is, every language has a Canonical Form that characterizes the surface properties and a standard semantic interpretation.

Above we pointed out the importance of a standard form in facilitating adult comprehension. There is no structural or architectural reason for this, rather Bever argues that it is true of attested languages because a language without it would not be learnable. This has some interesting implications for apparent structural universals — for example, Bever argues that the Extended Projection Principle (originally, that every sentence must have a subject) is actually the result of the pressure for a Canonical Form, and not a part of universal syntactic architecture.

### 8.6.  *Why Do We Think We Perceive Speech Almost Simultaneously with Its Acoustic Representation?*

Correspondingly, if an AxS scheme applies to vision, how does the derivational sequence of computations relate to the serial nature of eye-fixation snapshots at the input level? One possible answer (proposed by Townsend & Bever 2001) is that the derivational structure is computed only slightly behind the initial surface analysis. Thus, the two representations of meaning meld into one internal representation in a kind of dynamic inner 'motion'. Bever & Townsend suggested

that this may account for the classically noted perceptual salience of words in sentences — the sentence structure gives a kind of internal meta-contrast-like percept of a representation that explodes.

### 8.7.    What Is the Relation between AxS and Formal Properties of Grammar?

It is an intrinsic feature of an effective analysis by synthesis scheme, that it computes representations in two ways one based on the 'outer form' of sentences, one based on the 'inner form'. The first is based on some sort of 'direct perception', the second on computational recreation of representation that reflects a generative process. Recently, several authors have raised the old idea that this duality is characteristic of language in particular: Sentences are serial but also hierarchically structured. The obvious application to today's biolinguistics of this classical duality is its implications for how language is processed (as in Townsend & Bever 2001); but a less obvious implication for the computational architecture of grammars has been raised in several previous articles (Medeiros 2008, Piattelli-Palmarini & Uriagereka 2008). Medeiros argues that X-bar theory provides the essential self-combining 'molecule' of syntactic derivation and represents the best compromise between the need for a recursive self-replicating structure, and the need for a serial output: On his interpretation, X-bar theory results in the maximally efficient 'packing' of serial elements with the smallest number of abstract nodes in a hierarchy. An intriguing result of this compromise is that as the number of serial nodes increases linearly, the number of underlying nodes increases in the Fibonacci series. Piattelli-Palmarini & Uriagereka then note the general ubiquity of the Fibonnacci series in the hierarchical segmentation of many linguistic levels, including syllable structure, metrical forms and syntactic phases. They observe that Fibonacci series in general are the compromise result of opposing physical forces. They then cite Townsend & Bever (2001) as articulating the notion of 'two' routes to processing meaning as built into the AxS scheme: An initial one based on serial patterns, and a final one based on computational derivation. They suggest that the compromise between serial tractability and computational generativity may explain the existence of syntactic 'phases', which themselves cyclically build up in a Fibonnaci series, consonant with Medeiros's ideas.

The concept of phases is an interesting hypothesis, that specifies the orderly stages in which syntactic/lexical information is transferred to semantic representation of a sentence, as the computational structure is computed. In this way, it may ultimately be demonstrable that the duality of language reflected in how it is learned and processed, will also provide a deep explanation of some aspects of syntactic architecture itself.

### 8.8.    If Each 'Level' of Representation has its Own AxS Cycle, how are they Cascaded to Flow in Parallel? How does the Emerging Output of Each One Affect the Processing of the Other Levels?

To accomplish such matching, multi-time resolution processing seems like a promising approach. If the comprehension system operates at two principled

(physiologically constrained) rates, there will exist regular temporal windows in which to align the information coming from different levels of analysis. On one view, a faster cycle, roughly at the gamma rate (~25–50 Hz), will align with a slower, integrative theta (4–8 Hz) rate and possibly an even slower, 'phrasal' delta rate (<3Hz). While 'local', level-internal representations will be processed at the higher clock speed, integration across levels will be executed every 200 ms or so (theta rate), permitting the integration and alignment. Recent neurobiological data favors such a multi-time resolution approach (Poeppel 2003, Boemio *et al.* 2005, Giraud *et al.* 2007).

### 8.9.　*We Note, without Elaboration Further Questions for the Future*

What is the tolerance between the stored initial representation and the output of the synthetic component to count as 'similar enough'? If the synthesized match is 'abstract' how does that 'fill in' the missing acoustic or structural details? Why do we think that the phonetic-phoneme-syllable mapping is the 'first' stage of language understanding, either temporally or even logically? Is the AxS system relevant only for acquisition, after that everything is recomputed into over-learned templates? Can the three major subroutines of AxS be isolated using the tools of cognitive neuroscience? In particular, can (i) the initial (perhaps template-based) triggering of hypotheses, (ii) the derivation/synthesis from abstract representations, and (iii) the comparator stages be shown and manipulated to understand their internal architectures?

### References

Arnal, Luc, Benjamin Morillon, Christian Kell & Anne-Lise Giraud. 2009. Dual neural routing of visual facilitation in speech processing. *Journal of Neuroscience* 29, 13445–13453.

Bar, Moshe, Karim S. Kassam, Avniel Singh Ghuman, Jasmine Boshyan, Annette M. Schmid, Anders M. Dale, Matti Hamalainen, Ksenija Marinkovic, Daniel Schacter, Bruce Rosen & Eric Halgren. 2006. Top-down facilitation of visual recognition. *Proceedings of the National Academy of Science* 103, 449–454.

Bar, Moshe. 2007. The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences* 11, 280–289.

Al Bawab, Ziad, Bhisksha Raj & Richard Stern. 2008. Analysis by synthesis features for speech recognition. In *IEEE ICASSP 2008*, 4185–4188.

Bell, C. Gordon, Hiroya Fujisaki, John M. Heinz, Kenneth N. Stevens & Arthur S. House. 1961. Reduction of speech spectra by analysis-by-synthesis techniques. *Journal of the Acoustical Society of America* 33, 1725–1726.

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and Language Development*, 277–360. New York: Wiley and Sons.

Bever, Thomas G. 1992. The demons and the beast: Modular and nodular kinds

of knowledge. In Ronan G. Reilly & Noel E. Sharkey (eds.), *Connectionist Approaches to Natural Language Processing*. Hove: Lawrence Erlbaum.

Bever, Thomas G. 2009. The individual and universal in language. In Massimo Piattelli-Palmarini, Juan Uriagereka & Pello Salaburu (eds.), *Of Minds & Language: A Dialogue with Noam Chomsky in the Basque Country*, 278–295. Oxford: Oxford University Press.

Bloomfield, Leonard. 1914. *The Study of Language.* New York: Henry Holt and Co.

Boemio, Anthony, Stephen Fromm, Allen Braun & David Poeppel. 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience* 8, 389–395.

Boulenger, Veronique, Olaf Hauk & Friedemann Pulvermüller. 2008. Grasping Ideas with the Motor System: Semantic Somatotopy in Idiom Comprehension. *Cerebral Cortex* 19, 1905–1914.

Canolty, Ryan, Erik Edwards, Sarang Dalal, Maryam Soltani, Srikantan Nagarajan, Heidi Kirsch, Mitch Berger, Nicholas Barbaro & Robert Knight. 2006. High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628.

Charniak, Eugene. 1997. Statistical parsing with a context free grammar and word statistics. *Proceedings of the 14th National Conference on Artificial Intelligence (AAAI-97)*, 598–603.

Chickering, David, Dan Geiger & David Heckerman. 1994. Learning Bayesian networks is np-hard. Technical Report MSR-TR-94-17, Microsoft Research, November 1994.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

de Cordemoy, Geraud. 1684 [1972]. *Discours Physique de la Parole*. Delmar, NY: Scholars' Facsimiles and Reprints.

Crain, Stephen, Rosalind Thornton & Drew Khlentzos. 2008. The case of the missing generalizations. *Cognitive Linguistics* 20, 145–155.

D'Ausilio, Alessandro, Friedemann Pulvermüller, Paola S. Salmas, Ilaria Bufalari, Chiara Begliomini & Luciano Fadiga. 2009. The motor somatotopy of speech perception. *Current Biology* 19, 381–385.

Fei-Fei, Li & Pietro Perona. 2005. A Bayesian hierarchical model for learning natural scene categories. *IEEE CVPR*, 524–531.

Fodor, Jerry & Merrill Garrett. 1966. Some reflections on competence and performance. In John Lyons & Roger J. Wales (eds.), *Psycholinguistic Papers*, 135–179. Edinburgh: Edinburgh University Press.

Fiorentino, Robert & David Poeppel. 2007. Compound words and structure in the lexicon. *Language and Cognitive Processes* 12, 953–1000.

Galantucci, Bruno, Carol Fowler & Michael Turvey. 2006. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review* 13, 361–377.

Gennari, Silvia & David Poeppel. 2003. Processing correlates of lexical semantic complexity. *Cognition* 89, 27–41.

Giraud, Anne-Lise, Andreas Kleinschmidt, David Poeppel, Torben E. Lund, Richard Frackowiak & Helmut Laufs. 2007. Endogenous cortical rhythms

determine cerebral specialisation for speech perception and production. *Neuron* 56, 1127–1134.

Goodman, Kenneth. 1967. Reading: A psycholinguistic guess game. *Journal of the Reading Specialist* May, 126–135.

Guenther, Frank. 2006. Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders* 39, 350–365.

Halle, Morris & Kenneth Stevens. 1959. Analysis by synthesis. In W. Wathen-Dunn & L. E. Woods (eds.), *Proceeding of the Seminar on Speech Compression and Processing*, Vol. II, paper D7.

Halle, Morris & Kenneth Stevens. 1963. Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8, 155–159. [Reprinted in Halle, Morris. 2002. *From Memory to Speech and Back*. Berlin: Mouton de Gruyter.]

Harris, Zellig. 1970 *Papers in Structuralist and Transformational Linguistics*. Dordrecht: Reidel.

Hickok, Gregory. 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21, 1229–1243.

Hickok, Gregory & David Poeppel. 2007. The cortical organization of speech perception. *Nature Reviews Neuroscience* 8, 393–402.

Hochstein, Shaul & Merav Ahissar. 2002. A view from the top. *Neuron* 36, 791–804.

Huang, Xuedong, Alex Acero & Hsiao-Wuen Hon. 2001. *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development.* Upper Saddle River, NJ: Prentice Hall PTR.

von Humboldt, Wilhelm. 1836 [1988]. *On Language*. Cambridge: Cambridge University Press.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates, Acoustics Laboratory*. MIT, Technical Report No. 13. [Republished 1967, 7th edn., Cambridge, MA: MIT Press.]

James, William. 1890. *The Principles of Psychology*. New York: Dover Publications.

Kohler, Evelyne, Christian Keysers, M. Alessandra Umilta, Leonardo Fogassi, Vittorio Gallese & Giacomo Rizzolatti. 2002. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science* 297, 846–848.

Kveraga, Kestutis, Avniel S. Ghuman & Moshe Bar. 2007. Top-down predictions in the cognitive brain. *Brain and Cognition* 65, 145–168.

Ladefoged, Peter & Donald Broadbent. 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America* 29, 98–104.

Lappin, Shalom & Stuart M. Shieber. 2007. Machine learning theory and practice as a source of insight into universal grammar. *Journal of Linguistics* 43, 393–427.

Liberman, Alvin, Francis Cooper, Donald Shankweiler & Michael Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74, 431–461.

Lotto, Andrew & Lori Holt. 2006. Putting phonetic context effects into context: A

commentary on Fowler (2006). *Perception and Psychophysics* 68, 178–183.

Marantz, Alec. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22, 429–445.

Marslen-Wilson, William & Lorraine Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8, 1–71.

McClelland, Jay & Jeffrey Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18, 1–86.

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264, 746–747.

Medeiros, David. 2008. Optimal growth in phrase structure. *Biolinguistics* 2, 152–196.

Miller, George A. 1962. Some psychological studies of grammar. *American Psychologist* 17, 748–762.

Miller, George A., Eugene Galanter & Karl Pribram. 1960. *Plans and the Structure of Behavior*. New York: Holt, Rinehart and Winston.

Miller, George A. & Noam Chomsky. 1963. Finitary models of language users. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2, 419–491. New York: Wiley.

Morgan, Emily, Keller, Frank & Steedman, Mark. In press. A bottom-up parsing model of local coherence effects. *Proceedings of the 32nd Annual Cognitive Science Society Meeting.*

Neisser, Ulric. 1967. *Cognitive Psychology*. New York: Meredith.

Norris, Dennis & John M. McQueen. 2008. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* 115, 357–395.

Papeo, Liuba, Antonino Vallesi, Alessio Isaja & Raffaela Rumiati. 2009. Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PloS ONE* 4, 1–11.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34, 37–90.

Piattelli-Palmarini, Massimo & Juan Uriagereka. 2008. Still a bridge too far? Biolinguistic questions for grounding language on brains. *Physics of Life Reviews* 5, 207–224.

Poeppel, David. 2003. The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication* 41, 245–255.

Poeppel, David & David Embick. 2005. The relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics. Four Cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Poeppel, David & Gregory Hickok. 2004. Towards a new functional anatomy of language. *Cognition* 92, 1–12.

Poeppel, David & Philip Monahan. 2010. Feedforward and feedback in speech perception: Revisiting analysis-by-synthesis. *Language and Cognitive Processes*. DOI: 10.1080/01690965.2010.493301.

Pulvermüller, Friedemann, Martina Huss, Ferath Kherif, Fermin Moscoso del Prado Martin, Olaf Hauk & Yury Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences* 103, 7865–7870.

Pulvermüller, Friedemann & Luciano Fadiga. 2010. Active perception: sensori-

motor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11, 351–360.

Riezler, Stefan, Tracy H. King, Ronald M. Kaplan, Richard Crouch, John T. Maxwell III & Mark Johnson. 2002. Parsing the Wall Street Journal using a lexical-functional grammar and discriminative estimation techniques. *Proceedings of the 40ᵗʰ Annual Meeting of the Association for Computational Linguistics (ACL-02)*, 271–278.

Rizzolatti, Giacomo. 2005. The mirror neuron system and its function in humans. *Anatomy and Embryology* 210, 419–421.

Skipper, Jeremy, Virginie van Wassenhove, Howard Nusbaum & Steven Small. 2007. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex* 17, 2387–2399.

Stekelenburg, Jeroen J. & Jean Vroomen. 2007. Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience* 19, 1964–1973.

Stevens, Kenneth. 2002. Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America* 111, 1872–1891.

Titov, Ivan & James Henderson. 2007. Incremental Bayesian Networks for structure prediction. *Proceedings of the 24th International Conference on Machine Learning* (ICML 2007).

Townsend, David J. & Thomas G. Bever. 2001. *Sentence Comprehension*. Cambridge, MA: MIT Press.

Wagers, Matt & Colin Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45, 395–433.

van Wassenhove, Virginie, Kenneth Grant & David Poeppel. 2005. Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences* 102, 1181–1186.

Wilson, Stephen M., Ayse P. Saygin, Martin I. Sereno & Marco Iacoboni. 2004. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience* 7, 701–702.

Wundt, Wilhelm. 1900 *Die Sprache*. Leipzig: Engelmann.

Yuille, Alan & Dan Kersten. 2006. Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences* 10, 301–308.

Zweig, Eytan & Liina Pylkkänen. 2009. A visual M170 effect of morphological complexity. *Language and Cognitive Processes* 24, 412–439.

*Thomas G. Bever*                                       *David Poeppel*
*University of Arizona*                                  *New York University*
*Department of Cognitive Science*                       *Department of Psychology*
*302 Communication Building*                            *6 Washington Place*
*Tucson, AZ 85721*                                      *New York, NY 10003*
*USA*                                                   *USA*
*tgb@email.arizona.edu*                                 *david.poeppel@nyu.edu*

# Language Learning and Language Universals

## Jeffrey Lidz

This paper explores the role of learning in generative grammar, highlighting interactions between distributional patterns in the environment and the innate structure of the language faculty. Reviewing three case studies, it is shown how learners use their language faculties to leverage the environment, making inferences from distributions to grammars that would not be licensed in the absence of a richly structured hypothesis space.

## 1.    Introduction

What does it mean to learn? Within the cognitive sciences, learning is treated as the creation of a system of mental representations in response to a collection of experiences (Chomsky 1975, Gallistel 1990). The learning organism's task is to infer from data the system that produced that data. In the relatively simple case of learning a word, say 'dog', the learner's job is to collect observations about the use of that word to infer what people in their speech community intend when they say 'dog'. Learning has occurred when the learner knows what thoughts people are having (and intend him to have) when they say 'dog'. We can tell what meaning the learner has acquired for that word on the basis of whether it judges new objects to be dogs or non-dogs. In general, the learning organism's responses to new situations reveal the inferences made on the basis of experience. These inferences, in turn, reveal the properties of the mental representations that underlie learning and use.

The acquisition of syntax is parallel in this respect to the acquisition of words or any other cognitive structure. We assess the representations of the learner by examining how he responds to new sentences. Indeed, nearly every psychologically oriented discussion of syntax begins with these two observations: (i) that we can produce and understand sentences we have never heard before, and (ii) that of the sentences we have never heard before, we can recognize that some are possible but others are not. Just as we can categorize animals we have never seen before as dogs or non-dogs, we can categorize strings of words we have never heard before as sentences or non-sentences.[1] This ability implies a learner

---

[1]    There is no loss of generality if what the learner acquires is a system for assigning

that responds to experience not simply by memorization, but by mapping experience onto representations that make predictions about what other sentences are possible and impossible.

The leading idea of Chomsky's early discussions of learning (Chomsky 1959, 1965, 1975; cf. Lees 1957: 406ff) is that the representations built by the language learner imply a non-obvious metric of similarity between the experienced sentences and the possible but as yet unencountered sentences. The dimensions we use to judge new sentences as possible or impossible appear to be highly abstract and removed from experience. Because these particular dimensions are not the only imaginable dimensions that the learner might have used to construct linguistic representations, Chomsky's argument was that these dimensions (i.e. the dimensions of linguistic analysis) must be supplied not by experience, but by the innate endowment of the child.

Take, for example, the empty category principle (Chomsky 1981, 1986). If it is a true description of our knowledge of English syntax that non-pronominal empty categories must be properly governed, then the learning theory must supply the tools out of which this generalization can be identified or constructed.[2] One of the primary sources of linguistic nativism is the observation that the explanatory pieces of linguistic representation (like proper government and the primitives out of which it is built, for example, c-command, theta-marking, barrier, etc.) are in a vocabulary so far removed from experience that it becomes implausible that both this vocabulary and the complex relations built out of it are induced from experience. If it is further true that a generalization stated over this vocabulary (like the ECP) holds of the syntax of every language, then it becomes a reasonable hypothesis that this generalization reflects a universal feature of linguistic representation, a property that every language must exhibit as a consequence of biological design.

And here is the hypothesized connection between learning and universals. Because the universals reflect constraints on possible representations, learners simply do not consider representations outside of the space defined by these constraints. In the context of a learning theory, identifying the range and limits of possible languages is tantamount to identifying the immanent structure of the child's language acquisition device. A single piece of explanatory machinery would account for both the range of possible linguistic variation and the language learner's initial hypothesis space about what a language can be.

Critically, however, identifying the learner's initial hypothesis space is not equivalent to providing a model of how the learner maps the input onto the appropriate representations (Fodor 1966, Pinker 1984). Such a model requires a procedure for mapping experience (as it is experienced by the learner) onto the representations that generated that experience, i.e. the grammar of the language.

---

convergent derivations to strings of words, mappings from sound to meaning, or triples of sounds, meanings and derivations. The point is just that whatever system the learner acquires for representing the language must be such that it can deal with the unencountered sentences appropriately.

[2]   I take no stance on the status of the ECP in grammatical theory. The point is that if the theory posits some kind of grammatical knowledge, the learning theory must either be able to construct that knowledge out of some more basic primitives or else it is a primitive built into the learner directly.

This kind of learning model can be seen as a kind of analysis by synthesis, in which the language learner approximates the grammar that generated his experience (Halle & Stevens 1964, Townsend & Bever 2001).[3]

It is standardly held that having a highly restricted hypothesis space makes it possible for such a learning mechanism to successfully acquire a grammar that is compatible with the learner's experience and that without such restrictions, learning would be impossible (Chomsky 1975, Pinker 1984, Jackendoff 2002). In many respects, however, it has remained a promissory note to show how having a well-defined initial hypothesis space makes grammar induction possible in a way that not having an initial hypothesis space does not (see Wexler 1990 and Hyams 1994 for highly relevant discussion).

The failure to cash in this promissory note has led, in my view, to broad skepticism outside of generative linguistics of the benefit of a constrained initial hypothesis space. Despite the fact that deep insights about the range and limits of syntactic variation have been achieved through the methods of comparative syntax (e.g., Kayne 2000, Richards 2001, Baker 2005, *inter alia*), researchers in adjacent areas of cognitive science have been less impressed with the idea that abstracting out the universal formal properties of natural language from the study of individual grammars would lead to progress in explaining language acquisition.

This skepticism derives from several sources. First, it is not clear how these formal theories make contact with developmental data from children learning their first language. Second, the constantly growing sophistication of computational data-mining techniques seems to undercut the premise that the input does not contain the information relevant to building grammatical representations (Elman *et al.* 1996, Christiansen & Chater 1999, Klein & Manning 2004). Finally, a host of research showing that even young infants are sophisticated statistical learners seems to further raise the possibility that learners can extract more from the input that was assumed by standard arguments from the poverty of the stimulus (e.g., Saffran *et al.* 1996, Gomez & Gerken 2000)

Recent work in our laboratory is beginning to link formal theories of language universals with statistical approaches to language learning. This work explicitly examines the kinds of information that is available to a distributional learner and how this information is used in the course of language acquisition (Pearl 2007, Sneed 2007, Syrett 2007, Viau 2007, Pearl & Lidz 2009, Syrett & Lidz 2009, Takahashi 2009, Viau & Lidz 2009,). Our hope is that this research will both remove skepticism of nativist approaches to language acquisition generally and, more positively, show how the inferences that learners make from input distributions are constrained by antecedent knowledge of the universal features of linguistic representation and the range of possible linguistic variation. This program aims to make good on the promise that a constrained hypothesis space helps learners to use the input effectively in acquiring a particular language.

---

[3]    In the current paper, we emphasize the nature of the mapping between the input and the acquired grammar by considering the information that learners use and the conclusions that they can reasonable draw from that information. An important open issue in this context is the role of on-line parsing mechanisms in implementing these kinds of inferences in real time. See Baier & Lidz (2009) and Lidz *et al.* (2010) for discussion of the role of on-line algorithms in making inferences in acquisition.

## 2.    Sensitivity to Input Distributions

In the last 15 years, there has been a resurgence of interest in infants' sensitivity to statistical features of the input language. A wide range of studies with infants, children and adults have demonstrated their ability to track statistical features of an artificial language and to use these features to learn generalizations about those grammars (for reviews, see Gomez & Gerken 2000 and Saffran 2003). However, it is important to emphasize that sensitivity to the statistical features of the exposure language leaves open the question of how this sensitivity contributes to the acquisition of a grammar. In every theory of language acquisition, the learner must be sensitive to features of the environment. The fundamental issue in the domain of distributional learning is what kind of learning mechanism this sensitivity feeds into.

In a learning framework in which the child brings the space of possible grammars to bear on the acquisition of a particular language, sensitivity to statistical features of the environment functions as input to a selective learning mechanism. Such a mechanism, restricted by the child's innate endowment to representing only those relations that grammars can represent, provides an algorithm for selecting the appropriate representation of the input. Given linguistic experience plus the (probably infinite) set of possible grammars, the learning mechanism selects that grammar which provides the best description of the input from within that space (see, e.g., Miller & Chomsky 1963, Fodor 1966, Yang 2002, and Pearl 2007).

So, to best understand the role that the learner's sensitivity to the environment plays in language acquisition, it is important to identify the deductive consequences of this sensitivity. As noted in opening, we learn about the acquired representations by examining how the learner approaches new situations. Thus, by exploring the range of new situations that are taken by the child as compatible with their experience, we can infer something about the content of the acquired representations above and beyond the information contributed by experience.

## 3.    Deductive Consequences of Phrase Structure: Constraints on Movement

Every syntactic theory recognizes that sentences in a human language are not simply linear strings of words. Rather, words in a sentence are arranged in nested hierarchical structures (Chomsky 1957, Jackendoff 1977). These structures make it possible, for example, to derive multiple interpretations from a single string. Consider the well-worn example:

(1)    ancient history teacher

There are two possible meanings for this string. On one interpretation, this string refers to a very old person who teaches history. On the other, it refers to a teacher of ancient history. The ambiguity of the string supports the idea that a single string can be structured in multiple ways, as shown in (2).

(2)   a.      [ancient [history teacher]]
      b.      [[ancient history] teacher]

Constituent structure representations provide explanations for (at least) three kinds of facts. First, constituents provide the units of interpretation, as just seen. Second, the fact that each constituent comes from a category of similar constituents (e.g., NP, VP, etc.) makes it such that a single constituent type may be used multiple times within a sentence, as in (3):

(3)      [$_{IP}$ [$_{NP}$ the cat] [$_{VP}$ ate [$_{NP}$ the mouse]]]

Third, constituents provide the targets for grammatical operations such as movement and deletion:

(4)   a.      I miss [the mouse]$_i$ that the cat ate ___$_i$.
      b.      The cat ate the mouse before the dog did [$_{VP}$ ~~eat the mouse~~].

Thompson & Newport (2007) make a very interesting observation about phrase structure and its acquisition: Because the rules of grammar that delete and rearrange constituents make reference to structure, these rules leave a kind of statistical signature of the structure in the surface form of the language. The continued co-occurrence of certain categories and their consistent appearance and disappearance together ensures that the co-occurrence likelihood of elements from within a constituent is higher than the co-occurrence likelihood of elements from across constituent boundaries.

They go on to argue that this statistical footprint could be used by learners in the acquisition of phrase structure. And they show that adult learners are able to use this statistical footprint in assigning constituent structure to an artificial language. But again, showing that learners are sensitive to the statistical features of the environment does not yet provide information about the acquired representations. It is impressive that learners learned about the constituent structure of an artificial language given only statistical information about that structure. But this demonstration remains silent about the character of the acquired representations and the inferences that these representations license.

In order to determine whether the acquired representations have properties that derive from the structure of the learner, it is important to identify their deductive consequences. Do learners know things about constituent structure (even if this structure is acquired using statistical features of the environment) that are not evident in the statistics themselves?

In order to answer this question, Eri Takahashi and I constructed a miniature artificial grammar containing internally nested constituents. In addition, the grammar contained rules which allowed for the repetition of constituents of a certain type, the movement of certain constituents and substitution of certain constituents by pro-forms. We then created a corpus of sentences from this language in which these rules applied often enough to provide statistical evidence for the constituent boundaries. In other words, the language provided statistical cues to the internal structure of the sentences.

Our first question, using this artificial language, was whether adults and infants could acquire constituent structure using only statistical information. The language was presented in contexts that did not provide any referential information, so that no meaning could be assigned to any of the words. And, there was no prosodic or phonological information of any kind that could serve as a cue to the phrase structure. So, to the extent that learners could acquire the phrase structure, they would have to do so through the statistical features of the exposure. In order to test whether the learners acquired the phrase structure, we asked whether they could distinguish novel sentences containing either moved constituents or moved non-constituents. Since only constituents can move in natural languages, we reasoned that if learners could distinguish moved constituents from moved non-constituents, it must be because they had learned the constituent structure of the artificial language. We found that both adults, after 36 minutes of exposure, and 18-month-old infants, after only 2 minutes of exposure, were able to do so (Takahashi & Lidz 2007, Takahashi 2009). Thus, the statistical footprint of constituent structure is detectable by learners and is usable in the acquisition of phrase structure.

Now, the exposure provided to the learners in this experiment included sentences containing movement. Although the particular sentences tested were novel, they exhibited structures that had been evident during the initial exposure to the language. We thus went on to ask whether the inference that only constituents can move derives from the learner's exposure to movement rules which apply only to constituents or whether this inference derives from the child's antecedent knowledge about the nature of movement rules in natural language.

To ask this question, we created a new corpus of sentences from our artificial language. In this novel corpus we included sentences in which (i) certain constituents were repeated in a sentence, (ii) certain constituents were optionally absent from a sentence, and (iii) certain constituents were replaced by pro-forms. This combination of operations created a statistical signature of the phrase structure of the language such that it was possible to identify the constituent boundaries in the language. However, in this input corpus we included no examples of movement. This made it possible for us to identify the locus of the learner's knowledge that only constituents can move. If this knowledge derives from the learner's experience in seeing movement rules, then we would expect learners to be unable to distinguish moved constituents from moved non-constituents. On the other hand, if the learner brings knowledge about what kinds of movement operations are possible in natural language to the learning task, then we would expect learners to correctly distinguish moved constituents from moved non-constituents.

We found that both adults and 18-month-old infants displayed knowledge of the constraint that only constituents can move, even when their exposure to the artificial language contained no instances of movement whatsoever. Thus, we can conclude that some of what is acquired on the basis of statistical information is not itself reflected in the statistics. Since the learners in this experiment had seen no examples of movement, their knowledge of the constraint that only constituents can move could not have come from the exposure language but rather must have come from the learners themselves.

In sum, identifying the constituency of a language has consequences for novel sentences with structures never before encountered. These deductive consequences reveal the structure of the learner over and above any role of distributional learning. Distributional learning therefore functions as part of a process of mapping strings onto the grammar that generated them. But some properties of the identified grammar are contributed by the learner's antecedent knowledge of the class of possible grammars.

## 4.      The Deductive Consequences of (In)Definiteness: The Interpretation of Bare Plural Subjects

It is important to recognize that this kind of argument is not limited to learning artificial grammars. For example, Sneed (2007) made an argument of exactly this form in examining children's interpretations of indefinite NPs. She showed, first, that there is a distributional difference between indefinite and definite NPs that could be used in a process of categorizing NPs as either definite or indefinite, and second, that this categorization licenses inferences about interpretation that are not themselves supported in the input.

Because indefinites are generally used to introduce discourse referents and definites are generally used to identify existing discourse referents (Heim 1982, Kamp 1982), indefinites are significantly more likely to be used on the first mention than definites are. Similarly, because old information is more likely to occur earlier in a sentence than new information (Prince 1992), definites are significantly more likely than indefinites to occur in subject position.

Sneed showed that there is ample evidence of these asymmetries in speech to children learning English so that determiners can be accurately classified as either definite or indefinite simply by tracking the relative likelihood of their occurring in first mention contexts and by tracking the relative likelihood of their occurring in subject position. Sneed argued (among other things) that bare plurals (e.g., *dogs*) can be classified as indefinites using this procedure.

As is well-known, bare plurals do occur in subject position some of the time (though as Sneed showed, not nearly as often as they occur in object position). When they do occur in subject position, they are often ambiguous between a generic and an existential interpretation (Diesing 1992). Consider the sentence in (5) and its interpretations in (6):

(5)      Crocodiles live in the swamp.

(6)      a.      It is generally true of crocodiles that they live in the swamp.
         b.      There are some crocodiles that live in the swamp.

In examining a corpus of child-directed speech, Sneed (2007) found that bare plural subjects were uniformly used generically (i.e. the interpretation (6a)) and were never used existentially.

This observation thus raises the question of whether children know that bare plural subjects can be used with an existential interpretation despite the fact

that speech to children simply does not contain examples illustrating this fact. If learners simply acquired the distributional features of the exposure language, then we would expect them to learn that bare plural subjects are obligatorily interpreted generically. On the other hand, if the interpretive properties of bare plurals follow from their meaning in concert with their syntactic position (as, for example, in Diesing 1992), then we might expect that once children have identified an expression as indefinite, its interpretive profile follows automatically.

Indeed, Sneed found that 4-year-old children were equally able to interpret a bare plural subject existentially and generically, despite the fact that they were apparently never exposed to the existential interpretation of such expressions.

There are two important lessons to be drawn from this work. First, while it is certainly true that there are statistical cues to an NPs classification as definite or indefinite, these cues are informative only to the extent that they are antecedently connected to a representation. That is, it is only because of their interpretive properties that indefinites are relatively less likely to occur in subject position than definites are. The conclusion that an expression occurring less likely as a subject than an object is an indefinite is valid only if the learner is using this asymmetry to feed a decision process about preexisting categories. The learner could not use this asymmetry to draw an inference about (in)definiteness unless this asymmetry derived from a fundamental feature of the representation that predicted this asymmetry to exist.

Second, to the degree that learners use distributional information in language acquisition, it must be that such information is used only when there is a question of how to represent a given part of the language. If learners simply tried to reproduce the distribution that they were exposed to, then we would not have found that children allow bare plural subjects to be interpreted existentially. In other words, learners must not be trying to determine whether bare plurals can be interpreted existentially. If they were, then they should have drawn the inference that they cannot. Rather, learners must simply be trying to classify nominal expressions as definite or indefinite. Any additional interpretive properties follow as a matter of grammar, independent of the distribution of these interpretations in the input.

Again, the deductive consequences of distributional learning reveal the contribution that the learner makes to language acquisition. And again, evidence of distributional learning is not evidence against the learner having a highly constrained hypothesis space. Rather, in this case, as in the previous one, distributional learning can be seen as feeding a selective process by which learners use the data to identify the grammar that generated that data (see also Syrett & Lidz 2009, Pearl & Lidz 2009).

## 5.    Selective Learning in the Acquisition of Ditransitives

The inferences from surface form to grammatical representation can also be significantly less direct. Consider, for example, the range of ditransitive constructions in English, Spanish and Kannada.

(7)   *English*

    a.    John sent the book to Mary.

    b.    John sent Mary the book.

(8)   *Kannada*

    a.    Hari  rashmi-ge  pustaka-vannu  kalis-id-a.
           *Hari  rashmi-DAT book-ACC       send-PST-3SM*
           'Hari sent a book to Rashmi.'

    b.    Hari  rashmi-ge  pustaka-vannu  kalis-i-**koTT-a**.
           *Hari  rashmi-DAT book-ACC       send-PP-BEN-PST-3SM*
           'Hari sent a book to Rashmi.'

    c.    Hari  pustaka-vannu  rashmi-ge  kalis-id-a.
           *Hari  book-ACC       rashmi-DAT send-PST-3SM*
           'Hari sent a book to Rashmi.'

    d.    Hari  pustaka-vannu  rashmi-ge  kalis-i-**koTT-a**.
           *Hari  book-ACC       rashmi-DAT send-PP-BEN-PST-3SM*
           'Hari sent a book to Rashmi.'

(9)   *Spanish*

    a.    Carmen envió el  libro  a su    profesora.
           *Carmen sent   the book  to her   professor*

    b.    Carmen **le**    envió el  libro  a su    profesora.
           *Carmen **CL**   sent   the book to her   professor*

    c.    Carmen envió a su    profesora  el  libro.
           *Carmen sent   to her   professor   the book*

    d.    Carmen **le**   envió  a  su profesora el libro.
           *Carmen **CL** sent    to her professor   the book*

Whereas English has two surface forms for ditransitives, Kannada and Spanish have four. Viau & Lidz (2009), building on earlier work by Harley (2002), Bleam (2003), and Lidz & Williams (2005), argue that despite these surface differences, there is a coherent mapping of ditransitive structures across languages. In particular, languages make available two kinds of ditransitives: Those with the theme asymmetrically c-commanding the goal and those with the goal asymmetrically c-commanding the theme. In English, these correspond to the prepositional dative (7a) and the double object construction (7b), respectively, with word order functioning as a surface correlate of the syntactic structure. In Kannada and Spanish, however, word order is not an expression of the underlying configurational structure. Rather, the structure with the theme c-commanding the goal is the morphologically unmarked form whereas the structure with the goal c-commanding the theme is the morphologically marked form. In Kannada, this morphological form is realized through the verbal auxiliary *koDu* (8b,d). In Spanish it is realized through the dative pronominal clitic *le* (9b,d).

      The argument for this way of carving up the data comes from two kinds of

facts. First, the goal argument functions as a kind of possessor in those configu-
rations where it is argued that the goal c-commands the theme (cf. Oehrle 1976).
Second, patterns of binding from one argument into the other also support this
classification.

The basic pattern of judgments with a quantified dative argument in shown
in (10), where DAT indicates the indirect object, marked with dative case, ACC
indicates the direct object, marked with accusative case, and Q- indicates which
of these noun phrases contains a quantifier.

(10) a.     Q-DAT$_x$ ACC$_x$ BEN

Rashmi  pratiyobba hudugan-ige   avan-a     kudure-yannu tan-du-koTT-aLu.
*Rashmi   every        boy-DAT       3SM-GEN horse-ACC        return-PPL-BEN.PST-3SF*
'Rashmi returned every boy his horse.'

b.     Q-DAT$_x$ ACC$_x$ unaffixed

Rashmi  pratiyobba hudugan-ige   avan-a     kudure-yannu tan-d-aLu.
*Rashmi   every        boy-DAT       3SM-GEN horse.ACC        return-PST-3SF*
'Rashmi returned every boy his horse.'

c.     ACC$_x$ Q-DAT$_x$ BEN

Rashmi  avan-a       kudure-yannu   pratiyobba hudugan-ige tan-du-koTT-aLu.
*Rashmi   3SM-GEN horse-ACC         every        boy-DAT       return-PPL-BEN.PST-3SF*
'Rashmi returned his horse to every boy.'

d.     *ACC$_x$ Q-DAT$_x$ unaffixed

  * Rashmi  avan-a       kudure-yannu   pratiyobba hudugan-ige   tan-d-aLu.
    *Rashmi   3SM-GEN horse-ACC         every        boy-DAT         return-PST-3SF*
'Rashmi returned his horse to every boy.'

Descriptively speaking, when the dative-marked object comes first (10a–b), it can
bind into the accusative-marked object, whether or not the benefactive affix is
present. In contrast, when the accusative-marked object comes first (10c–d), the
dative can bind into it only in the presence of the benefactive affix.

If the quantificational phrase is the accusative argument and the prono-
minal is contained in the dative argument, however, a different pattern emerges.

(11) a.     *DAT$_x$ Q-ACC$_x$ BEN

  * Sampaadaka adar-a     lekhan-ige   pratiyondu lekhana-vannu kaLis-i-koTT-a.
    *editor          it.GEN    author.DAT every         article.ACC        send-PP-BEN.PST-3SM*
'The editor sent its author every article.'

b.     DAT$_x$ Q-ACC$_x$ unaffixed

Sampaadaka adar-a     lekhan-ige   pratiyondu lekhana-vannu   kaLis-id-a.
*editor          it-GEN    author-DAT every         article-ACC        send-PST-3SM*
'The editor sent its author every article.'

    c.       Q-ACC$_x$ DAT$_x$ BEN

| Sampaadaka | pratiyondu | lekhana-vannu | adara | lekhan-ige | kaLis-i-koTT-a. |
|---|---|---|---|---|---|
| *editor* | *every* | *article-ACC* | *it-GEN* | *author-DAT* | *send-PP-BEN.PST-3SM* |

        'The editor sent every article to its author'.

    d.       Q-ACC$_x$ DAT$_x$ unaffixed

| Sampaadaka | pratiyondu | lekhana-vannu | adar-a | lekhan-ige | kaLis-id-a. |
|---|---|---|---|---|---|
| *editor* | *every* | *article-ACC* | *it-GEN* | *author-DAT* | *send-PST-3SM* |

        'The editor sent every article to its author'.

Here we see that when the accusative-marked object comes first (11c–d), it can bind into the dative-marked object, regardless of whether the benefactive affix is present on the verb. However, when the dative-marked object comes first (11a–b), the accusative-marked object can bind into it only when the benefactive affix is absent. The relevant binding possibilities for quantified dative and accusative arguments are summarized below.

(12)  a.   √ Q-DAT$_x$ ACC$_x$ V-**BEN**      e.   * DAT$_x$ Q-ACC$_x$ V-**BEN**
       b.   √ Q-DAT$_x$ ACC$_x$ V           f.   √ DAT$_x$ Q-ACC$_x$ V
       c.   √ ACC$_x$ Q-DAT$_x$ V-**BEN**      g.   √ Q-ACC$_x$ DAT$_x$ V-**BEN**
       d.   * ACC$_x$ Q-DAT$_x$ V          h.   √ Q-ACC$_x$ DAT$_x$ V

     Lidz & Williams (2005) argue that the above asymmetries arise from there being two distinct underlying structures for ditransitives in Kannada. When the benefactive affix is present, the DAT–ACC order is the underlying order, with the ACC–DAT order derived by A-movement. Thus, according to Lidz & Williams (2005), (12c) is derived from (12a) by movement of the accusative argument past the dative. The appearance of backward binding in (12c) is due to the fact that A-movement of the accusative over the dative does not destroy the binding relation established in the underlying order (12a), in which the quantified dative NP c-commands and thereby grammatically binds into the accusative. Similarly, since the DAT–ACC order is underlying, a quantificational accusative-marked object cannot bind into the dative (12e) unless A-movement has occurred, introducing a new configuration to license binding (12g).
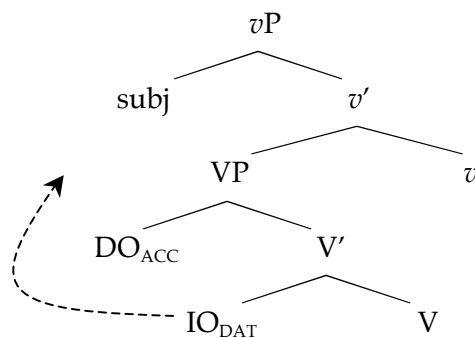
     When there is no benefactive affix, the ACC–DAT order reflects the underlying structure, and the DAT–ACC order is derived by A-movement. Thus, (12b) is derived from (12d). The quantified dative NP in (12d) cannot bind into the accusative because it does not c-command the accusative; only after moving above the accusative, as in (12b), can it grammatically bind into the accusative. By the same logic, since the accusative is underlyingly higher than the dative, the binding of the dative by the accusative can be established over this represen-tation (12h) and subsequent A-movement will not destroy it (12f).

     We will assume this analysis as well as the syntactic representations that it entails, shown in (13) with the benefactive affix and in (14) without. Optional A-movement is marked with a dashed arrow.

(13)  Subject DAT ACC V-BEN

$$vP$$

```
              vP
          /        \
       subj          v'
               /          \
            BENP            v
          /      \
      IO_DAT     BEN'
               /       \
             VP         BEN
           /    \
       DO_ACC    V
```

(14)  Subject ACC DAT V

$$vP$$
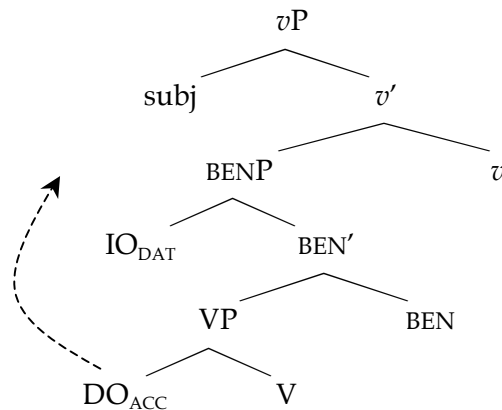
```
              vP
          /        \
       subj          v'
               /          \
             VP             v
           /     \
       DO_ACC    V'
                /    \
            IO_DAT    V
```

The conclusion that there are two distinct underlying structures in Kannada, each of which can be transformed by A-movement of the lower NP past the higher one straightforwardly captures the binding asymmetries discussed above.

The variable binding facts just presented have exact analogs in Spanish (Bleam 2003) with the clitic doubled ditransitives functioning exactly as the benefactive ditransitives in Kannada, and the morphologically unmarked ditransitives functioning exactly alike in the two languages.

In sum, so far, English, Kannada, and Spanish all utilize essentially the same two structures in ditransitives. In the 'prepositional dative', the Accusative argument c-commands the Dative underlyingly. This structure is expressed in English as the prepositional dative, in Spanish as the non-clitic-doubled ditransitive, and in Kannada as the non-benefactive ditransitive. In the 'DO-Dative', the Dative argument c-commands the Accusative underlyingly. This structure is expressed in English as the DOC, in Spanish through dative clitic doubling, and in Kannada through the benefactive verbal affix. Importantly, despite this fundamental structural symmetry across the three languages, the surface manifestation of these structures is distinct in each language. DO-datives show a distinct surface word order in English but not in Kannada and Spanish. Both Kannada and Spanish, unlike English, have a morphological distinction which correlates with the choice of DO-dative or prepositional dative. In Kannada, the DO-dative variant is marked by a benefactive verbal affix, while in Spanish this variant is marked via clitic doubling of the dative argument.

Given this characterization of the facts, learners cannot rely either on the word order or the morphological form as evidence in determining which of the two abstract structures underlies a ditransitive sentence in the language they are learning. Neither cue is cross-linguistically reliable (Haspelmath 2005). Thus, to the extent that learners can be shown to identify the appropriate structure we are faced with an interesting puzzle. The fact that the very same structures are exhibited in languages with such divergent surface syntax points towards just the sort of cross-linguistic commonality that a selective learning theory is intended to explain, since the pieces of explanation are identical across languages. On the other hand, the fact that the surface realizations of these structures diverge across languages would appear to make it difficult to use the surface form as a cue for the underlying structure (see Hyams 1986 and Snyder 1995 for related problems in other syntactic domains).

Viau & Lidz (2009) demonstrate first that four year old learners of Kannada already command the variable binding facts just described, placing an upper bound on how much time it takes learners to acquire these facts. Moreover, it seems highly unlikely that learners could have acquired the full range of binding possibilities simply by being exposed to positive examples of the relevant sort. The kinds of sentences that exhibit these asymmetries are exceedingly rare. And, even if these sentences did occur, there is no guarantee that the learner would know what the intended interpretation was or whether other interpretations were possible but simply not yet encountered. Thus, Viau & Lidz looked for a more indirect source of evidence that learners could use to identify the correct representation. We argued specifically that the relevant source of evidence resides in the distribution of animate and inanimate goal arguments.

The selective learning account works as follows. The child comes to the learning task with the knowledge that natural languages use at least two ways to configure ditransitives: A possession-based structure in which the IO occurs higher than the DO, and a location-based structure in which the DO occurs higher than the IO. If the child is faced with two distinct types of ditransitive clauses (e.g., DO- vs. prepositional dative in English, benefactive vs. non-benefactive in Kannada, clitic-doubled vs. not clitic doubled in Spanish), she must then identify which of these to associate with which underlying configuration. To do so, the child relies on the distribution of animate IOs. The construction in which IOs are more likely to be animate than inanimate has the possession configuration, since inanimates are highly unlikely possessors. And the construction in which IOs are more likely to be inanimate than animate has the location configuration, for the same reason.

Importantly, once the learner correctly identifies the underlying configuration, the variable binding asymmetries that we have observed in our experimentation follow directly. Thus, the learner requires no experience with particular binding configurations in order to acquire the variable-binding asymmetries we have observed in our experimentation.[4]

---

[4]   This reasoning is exactly parallel to the discussion of the 'compounding parameter' in Snyder 1995. The learner can observe productive root compounding as a surface correlate of a particular structure from which many apparently unrelated facts follow.

This is not to say, of course, that there is no learning involved in the acquisition of ditransitives. Our account is a learning-theoretic account in which the child, armed with a set of possible configurations for ditransitives and faced with the data, is able to use certain patterns of distribution to identify a mapping between surface forms and innate configurations. The innate guidance comes from the set of configurations and their semantic properties. Knowing these semantic properties enables the learner to track appropriate distributional information in the surface forms in order to learn which surface forms map onto which of the innate configurations (cf. Hyams 1986). Again, the configurations, in concert with basic structural requirements on variable binding and knowledge of how word order can be manipulated in the target language (which surely is at least partially learned), directly determine the binding possibilities for ditransitives.

The critical feature of this account in the current context is that the asymmetry in animacy of dative arguments across constructions in a useful cue to the underlying structure only if the learner comes equipped with knowledge of the class of possible ditransitive structures. Without that knowledge, the asymmetry in animacy is completely uninformative about the hierarchical structure of the clause.

## 6.    Conclusion

In this paper, I have sketched some results pointing towards an integrated theory of language acquisition that strongly bridges theoretical work on grammatical structure with developmental work on first language acquisition. It has long been held that the theory of language universals is equivalent to the theory of the language learner. This equivalence is only partial, however, since it only relates language universals to a specification of the initial state of the language learner. It is silent with respect to the algorithms that learners use to map their experience onto particular grammars within the space defined by universal grammar. The work described in the current paper shows how it is possible to find evidence about the structure of the learner's initial state from developmental data revealing how learners use distributional, statistical facts about the language they experience to acquire a particular grammar.

In addition, the work I have described demonstrates the convergence between statistical approaches to language learning with traditional nativist approaches. The fact that learners are highly sophisticated when it comes to identifying statistical regularities in the environment does not by itself provide evidence either for or against a learning mechanism driven by innate knowledge of the space of possible representations. It is only when we identify the deductive consequences of statistical learning that we begin to see how statistical learning works in the service of grammatical inference. In the cases reviewed here, the deductive consequences of statistical learning are very rich, pointing to a highly articulated hypothesis space over which statistical inference can be carried out in language acquisition.

## References

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Baier, Rebecca & Jeffrey Lidz. 2009. Predictive parsing impedes word learning in 19-month-olds. Poster presented at the *34ᵗʰ Boston University Conference on Language Development (BUCLD 34)*, Boston. (Boston University, 6–8 November 2009)

Bleam, Tonia. 2003. Properties of the double object construction in Spanish. In Rafael Núñez-Cedeño, Luis López & Richard Cameron (eds.), *A Romance Perspective on Language Knowledge and Use,* 233–252. Amsterdam: John Benjamins.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1959. A Review of B. F. Skinner's Verbal Behavior. *Language* 35, 26–58.

Chomsky, Noam. 1975. *Reflections on Language*. New York: Pantheon.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Christiansen, Morten & Nick Chater. 1999. Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23, 157–205.

Diesing, Molly. 1992. *Indefinites.* Cambridge, MA: MIT Press.

Elman, Jeffrey, Elizabeth Bates, Mark H. Johnson & Domenico Parisi. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Fodor, Jerry A. 1966. How to learn to talk: Some simple ways. In Frank Smith & George Miller (eds.), *The Genesis of Language*, 105–122. Cambridge, MA: MIT Press.

Gallistel, Charles R. 1990. *The Organization of Learning*. Cambridge, MA: MIT Press.

Gomez, Rebecca & LouAnn Gerken. 2000. Infant artificial language learning and language acquisition. *Trends in Cognitive Science* 4, 178–186.

Halle, Morris & Kenneth Stevens. 1964. Speech recognition: A model and program for research. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs, NJ: Prentice Hall.

Harley, Heidi. 2002. Possession and the double object construction. In Pierre Pica & Johan Rooryck (eds.), *The Yearbook of Linguistic Variation* 2, 29–68. Amsterdam: John Benjamins.

Haspelmath, Martin. 2005. Ditransitive constructions: The verb 'give'. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 426–429.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Amherst, MA: University of Massachusetts dissertation.

Hyams, Nina. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.

Hyams, Nina. 1994. Discreteness and variation in child language: Implications for principle and parameter models of language development. In Yonata Levy

(ed.), *Other Children, Other Languages*, 11–40. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jackendoff, Ray. 1977. *X-bar Syntax*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of Language*. Oxford: Oxford University Press.

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal Methods in the Study of Language*, 277–322. Amsterdam: Mathematisch Centrum.

Kayne, Richard S. 2000. *Parameters and Universals*. New York: Oxford University Press.

Klein, Dan & Christopher D. Manning. 2004. Corpus-based induction of syntactic structure: Models of dependency and constituency. *Proceedings of the 42$^{nd}$ Annual Meeting of the ACL*, 478–485.

Lees, Robert. 1957. Review of Chomsky's Syntactic Structures. *Language* 33, 375–408.

Lidz, Jeffrey, Ann Bunger, Erin Leddon, Rebecca Baier & Sandra Waxman. 2010. When one cue is better than two: Syntactic vs. lexical cues in infant verb learning. Unpublished manuscript, University of Maryland, College Park.

Lidz, Jeffrey & Alexander Williams. 2005. C-locality and the interaction of reflexives and ditransitives. In Keir Moulton & Matthew Wolf (eds.), *Proceedings of NELS 34*, vol. 2, 389–404. New York: Stony Brook University.

Miller, George & Noam Chomsky. 1963. Finitary models of language users. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. II, 268–321. New York: Wiley.

Oehrle, Richard. 1976. The grammatical status of the English dative alternation. Cambridge, MA: MIT dissertation.

Pearl, Lisa. 2007. Necessary bias in language learning. College Park, MD: University of Maryland dissertation.

Pearl, Lisa & Jeffrey Lidz. 2009. When domain general learning succeeds and when it fails. *Language Learning and Development* 5, 235–265.

Pinker, Steven. 1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Prince, Ellen. 1992. The ZPG letter: Subjects, definiteness, and information-status. In Sandra Thompson & William Mann (eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*, 295–325. Amsterdam: John Benjamins.

Richards, Norvin. 2001. *Movement in Language: Interactions and Architectures*. Cambridge: Cambridge University Press.

Saffran, Jenny. 2003. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science* 12, 110–114.

Saffran, Jenny, Richard Aslin & Elissa Newport. 1996. Statistical learning by eight-month-old infants. *Science* 274, 1926–1928.

Sneed, Elisa. 2007. The acquisition of generics. Evanston, IL: Northwestern University dissertation.

Snyder, William. 1995 Language acquisition and language variation: The role of morphology. Cambridge, MA: MIT dissertation.

Syrett, Kristen. 2007. Learning about the structure of scales: Adverbial modification and the acquisition of the semantics of gradable adjectives.

Evanston, IL: Northwestern University dissertation.

Syrett, Kristen & Jeffrey Lidz. 2009. 30-month-olds use the distribution and meaning of adverbs to interpret novel adjectives. Unpublished manuscript, Rutgers University, New Brunswick, NJ & University of Maryland, College Park.

Takahashi, Eri. 2009. Beyond statistical learning in the acquisition of phrase structure. College Park, MD: University of Maryland dissertation.

Takahashi, Eri & Jeffrey Lidz. 2008. Beyond statistical learning in syntax. In Anna Gavarró & M. João Freitas (eds.), *Proceedings of Generative Approaches to Language Acquisition (GALA) 2007*, 446–456. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Thompson, Sandra & Elissa Newport. 2007. Statistical learning of syntax: The role of transitional probability. *Language Learning and Development* 3, 1–42.

Townsend, David & Thomas Bever. 2001. *Sentence Comprehension: The Integration of Habits and Rules*. Cambridge, MA: MIT Press.

Viau, Joshua. 2007. Possession and spatial motion in the acquisition of ditransitives. Evanston, IL: Northwestern University dissertation.

Viau, Joshua & Jeffrey Lidz. 2009. Selective learning in the acquisition of Kannada ditransitives. Unpublished manuscript, Johns Hopkins University, Baltimore, MD & University of Maryland, College Park.

Wexler, Kenneth. 1990. Innateness and maturation in linguistic development. *Developmental Psychobiology* 23, 645–660.

Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

*Jeffrey Lidz*
*University of Maryland*
*Department of Linguistics*
*1401 Marie Mount Hall*
*College Park, MD 20742*
*USA*
*jlidz@umd.edu*

# Two Case Studies in Phonological Universals:
# A View from Artificial Grammars

## Andrew Nevins

This article summarizes the results of two experiments that use artificial grammar learning in order to test proposed phonological universals. The first universal involves limits on precedence-modification in phonological representations, drawn from a typology of ludlings (language games). It is found that certain unattested precedence-modifying operations in ludlings are also dispreferred in learning in experimental studies, suggesting that the typological gap reflects a principled and universal aspect of language structure. The second universal involves differences between vowels and consonants, and in particular, the fact that phonological typology finds vowel repetition and harmony to be widespread, while consonants are more likely to dissimilate. An artificial grammar task replicates this bias in the laboratory, suggesting that its presence in natural languages is not due to historical accident but to cognitive constraints on the form of linguistic grammars.

## 1.    Introduction: Phonological Universals and Artificial Grammars

When asked, "So you are a linguist — how many languages do you speak?", a tongue-in-cheek way that I often respond is "No, that's not what the point of generative linguistics is — what matters is how many impossible languages I don't (and could never) speak". A commitment to biologically-based universal preferences for certain types of grammatical structures over others makes clear predictions about what occurs when individuals are confronted with what Moro (2008) calls "impossible languages" — languages that contain structures not derivable from the primitives of Universal Grammar. Among these predictions, one is that individuals attempting to acquire an impossible language through the manner that languages are naturally acquired by children will not fully master an impossible pattern; see Smith & Tsimpli (1995) for a suggestive case study. A second prediction is that, should structures be introduced into a language that are not compatible with aspects of the universal blueprint/template for natural language, they will not remain stable across generations of users; see Kegl *et al.*

(1999) for discussion of a relevant case study, and, with a different methodology, Kirby *et al.* (2008). One last prediction is that an attempt to learn and use such patterns would not be represented in the same neural circuits that mediate natural languages; see Musso *et al.* (2003) for a revealing experiment of this sort. In short, the study of impossible languages and their acquisition, their non-persistence over generations, and their neural representation can be highly revealing to aspects of biological universals of language and to their feasibility across multiple timescales.

In this article, I will discuss two universals in the organization of phonological systems: The privileged position of edges in intersyllabic processes, and the asymmetric roles of consonants and vowels in intersegmental processes. My goal is to illustrate the study of universals at two different levels of phonological structure in a more general light by looking closely at different types of case studies. Both will be informed by experiments that use the artificial grammar methodology in order to investigate relative ease of learnability and generalizability of unattested grammatical patterns. By creating artificial and controlled examples of these unattested patterns we can observe whether they are unattested because of pure historico-geographic accident or due to more principled reasons, such as Universal Grammar — a set of analytic biases that prefer certain language types over others. Indeed, Ohala (1986), in his 'Consumer's guide to phonological evidence', recommends invented language games as among the best types of evidence for phonological representations and processes. While typological and theoretical research often repeatedly uncovers a number of universals, and this has the business of various schools of linguistics from Greenberg (1963) to the Principles-and-Parameters framework of Chomsky (1981) to the Optimality Theory model of Prince & Smolensky (1993), sometimes it only takes a few skeptics to say that we simply haven't found enough languages to know whether this is a true generalization or not, and that perhaps waiting for us in the Amazon is a language that violates exactly the universal we take to be central to human language structure.

It is my contention that one of the most effective ways of examining whether there is a true analytic and cognitive bias for one type of linguistic structure over another is in teaching it to experimental participants who have neither in their native language, and seeing whether they learn or prefer one to the other. This pursuit is reminiscent of Hauser's (2009: 190) question:

> Do animal forms fill up the space of possible forms or, more generally, does the genome have the potential to create an unbounded range of variation with no gaps? Answers to this question are only beginning to emerge, but they suggest that there are at least three factors that constrain the range of potential forms, creating gaps that have never been, and may never be, filled.

In other words, certain morphological structures in organisms are unattested, not only as a result of "sampling error" due to a paucity of earth-scouring specimen collection, but because of various factors that Hauser identifies as rendering certain organismal forms impossible: Phylogenetic inertia, lack of relevant environmental pressures that result in selection among the biologically given

options, and physical design constraints. In short, Hauser's conclusion is that, due to the interplay of biologically-determined primitives of organismic form, conditions of generational change, and environmental pressures, certain *logically possible* forms that have never arisen may be *bio*logically *impossible*, and thus may never arise (see also Boeckx & Piattelli-Palmarini 2005: 449). In the domain of language, we find analogous factors at work in explaining why certain linguistic structures are never found: Persistence of successful or efficient linguistic structures, absence of contact or relevant noise that would lead to reanalysis/parameter setting, and certain constraints on what language must deliver to the articulatory-perceptual and conceptual-intentional interfaces.

In what follows, I address two case studies in the organization of phonological systems based on typological research that suggest certain structures are impossible, and use the experimental methodology of artificial grammar learning to test whether such impossible languages can be used and acquired as easily as closely-matched but linguistically natural patterns.

## 2.    Universals of Precedence-Modifying Ludlings

Our first example, related to the abstract representations of precedence among syllables within a word, comes from what at first may appear to be an unusual domain of language use. Language games (think of Pig Latin, for example) — or *ludlings*, as they have been called by Laycock (1972) — exist in virtually every culture, usually among adolescents, either for the social function of group membership ('secret handshakes') or in order to encode/hide information from one's parents/rivals. While ludlings fall into many types, including iterative infixation (e.g., English *ubbi-dubbi*, Spanish *Jerigonza*, Portuguese *Língua do Pê*), perhaps the best known type are precedence-modifying ludlings that operate at the level of syllables, of which French *Verlan* (from *à l'envers*) is most famous (Plenat 1995). Syllable-precedence-modifying ludlings exchange the order of syllables in a word and are most commonly employed in disyllabic words; for example, Verlan transforms French *barjot* 'crazy' into → *jobard*.

Bruce Bagemihl, one of the most ardent proponents of ludlings as an object of linguistic study and as a source of information about possible and impossible operations in the phonological component, conducted an extensive typology of attested and non-attested ludlings (Bagemihl 1989). Some of Bagemihl's generalizations are listed below.

(1)    i.      No ludling reverses the middle two syllables.
              (e.g., *bar.go.tu.li → bar.tu.go.li*)

       ii.     No ludling moves the final syllable to the arithmetic middle.
              (e.g., *bar.go.tu.li.na → bar.go.na.tu.li*)

       iii.    No ludling permutes every other segment in a word.
              (e.g., *bram.poj → am.brjop*)

       iv.     No ludling permutes feet.
              (e.g., *bar.go.tu.li → tu.li.bar.go*)

> v.    No ludling permutes sub-segmental features
>        (e.g., *tom.duk → nob.tug*)

Following Bagemihl's insight that "ludlings extend, modify, or exaggerate attested natural language processes" (p. 492), we concur that precedence-modifying ludlings constitute a rich source of information about spontaneous transformations on phonological representations, free of prescriptive influence, and that given the wide variety of ludling processes, it can be quite revealing what one *doesn't* find.

Perhaps one of the more interesting findings about ludlings in the world at large is the fact that, while disyllabic reversals of the Verlan type are extremely common, one encounters a great deal of variation with words of longer syllable-counts. An immediate question that arises is the source of this variation: Is anything possible? We submit that this variation emerges as the consequence of ambiguity as to the way of representing the basic transformation in disyllabic forms. Indeed, there are at least five different attested ways of performing precedence-modification on words longer than two syllables:

(2)    Ambiguity of disyllabic inversion leads to variation on longer forms:
    a.    *pii.roo.wal → roo.wal.pii*        (Move σ1 (first) to end)    *Fula*
    b.    *ka.ma.tis → tis.ka.ma*            (Move σF (final) to start)    *Tagalog*
    c.    *nu.ku.hi.va → ku.nu.hi.va*        Transpose(σ1, σ2)          *Marquesan*
    d.    *ya.mu.nu.kwe → ya.mu.kwe.nu*    Transpose(σF, σF-1)        *Luchazi*
    e.    *va.li.si → si.li.va*                Invert order of all σ        *Saramaccan*

What is highly interesting about the five patterns in (2) is the fact that *all of them* are compatible with the disyllabic pattern σ1 σ2 → σ2 σ1. That is, σ1 σ2 → σ2 σ1 can indeed be analyzed as movement of σ1 to the end (2a), movement of σF to the beginning (2b), transposition of σ1 and its immediate successor (2c), transposition of σF and the immediately preceding syllable (2d), or total inversion of the order (2e). It is indeed plausible to think that all five patterns in (2) represent *different ways of generalizing* from the *same ambiguous input*. These ways of extending the disyllabic pattern to tri- and tetra-syllabic patterns have the potential to inform us about how learners generalize based on limited input. However, in the case of ludlings, we do not always know the full corpus of input data, nor whether learners are 'explicitly trained' on how to play, and whether they receive negative evidence or corrections.

One of the best ways to investigate 'poverty of the stimulus' type questions — that is, the question of how learners generalize a pattern from limited input to rarer or differing environments for application — is when the researcher has the ability to control *exactly how poor* the stimulus is. To this end, we decided to conduct an experiment in which we taught a ludling to volunteer participants, controlling exactly what kind of data they would be learning from in the training session prior to testing for generalization.

In Nevins & Endress (2007), we conducted an experiment in which participants were presented with an ambiguous rule involving trisyllabic sequences of nonce syllables: 123 → 321 (e.g., *ka.lei.bo → bo.lei.ka*). This

transformation is compatible with at least four hypotheses:

(3)    i.      Invert the order of syllables.
       ii.     Exchange the first and last syllable.
       iii.    Exchange the final and antepenultimate syllable.
       iv.     Exchange every other syllable (i.e. σj with σj+2).

These hypotheses differ in the instances or kinds of positions they explicitly name, for example, first, last, antepenult. In principle, upon hearing 123 → 321, participants might have chosen any of the hypotheses in (3), all of which account the data. Importantly, these four hypotheses all diverge on their predictions for an input string in which there are tetrasyllabic inputs, as shown for the hypotheses in (3) in their respective order:

(4)    i.      Invert the order of syllables: 1234 → 4321
       ii.     Exchange the first and last syllable: 1234 → 4231
       iii.    Exchange the final and antepenultimate syllable: 1234 → 1432
       iv.     Exchange every other syllable (i.e. σj with σj+2): 1234 → 3412

The hypotheses in (3iii) and (3iv) are unexpected based on the existing typology of ludlings. There are no extant precedence-modifying ludlings that refer to 'penultimate' or 'every other' syllable. There are two ways to interpret this typological lacuna. One is the result of a sampling error, the failure to find such a ludling due to not looking enough or having too small of a sample size in the world's languages. The other is that it represents a principled gap that is the result of an analytic bias (e.g., Universal Grammar), namely, that 'penultimate' or 'every other' syllable are predicates that are disfavored or disallowed in the construction of hypotheses that generalize to strings of different lengths. On the other hand, (3i) and (3ii) are not only attested in surveys of precedence-modifying ludlings, they are built on primitives that recur time and again in linguistic structural descriptions. We turn briefly to a discussion of the importance of the predicates 'first' and 'last' syllable within the more general context of 'edges of sequences'.

Starting with Ebbinghaus (1885/1913), it has been acknowledged that not all positions in sequences behave in the same way: Items close to the sequence's edges (that is, in the first and the last position) seem to be remembered better than items in other positions. This effect, however, seems to have different sub-components. Learners do not only remember *that* an item occurred in a sequence, but also *where* in the sequence it occurred; that is, they memorize also the *positions* of items. The memory for positions is most impressively illustrated by intrusion errors in memorization experiments (e.g., Conrad 1960). In such mistakes, participants erroneously recall elements from another list than the one currently recalled; these intrusions, however, often respect the positions in which they occurred in their original list. It thus seems that participants memorize an item's abstract sequential position (e.g., Hicks *et al.* 1966, Schulz 1955). This and related research has revealed that also the positions of items (and not only the identity of items themselves) are remembered better in *edges* than in other positions;

accordingly, most recent models of memory for positions in sequences assume, in some form or another, that only edges have absolute positional codes, and that internal positions are encoded relative to the sequences' edges (e.g., Henson 1998, Hitch *et al.* 1996, Ng & Maybery 2002).

Taking these results as the foundation for constraints on linguistic primitives, we suggest that two-argument operations of precedence-modifying ludlings of transposition (indicated below by [x><y] to transpose x and y (Halle 2008)) can only occur with an edge-syllable and with a syllable defined by a function relativized to that edge:

(5)     Repertoire of allowed precedence-modifying operations in natural language:
   i.      Total inversion
   ii.     Transpose operations limited to *x,f(x)*, where *x* can be FIRST, LAST
              and where *f(x)* can be
           PRECEDER(x): The element immediately preceding *x*      (e.g., 1[2><3])
           SUCCEEDER(x): The element immediately following *x*      (e.g., [1><2]3)
           COMPLEMENT(x): The entire sequence in the word excluding *x*
                                                  (e.g., [1><23] or [12><3])
           POLAR(x):   The opposite edge of the word from *x*      (e.g., [1>234<5])
           DOPPEL(x):  The corresponding position to *x* in an adjacent word
                                                  (e.g., [1>234] [<7]89)

Some examples of the uses of transposition operations on these functions from existing ludlings are shown in (5), where these are typed functions that can occur not only over syllables, but also sub-syllabic constituents such as onset, nucleus, and body (onset plus nucleus):

(6)   a.    *dito → doti*                                *Tagalog*
            Transpose (first, successor) over Nucleus

      b.    *wudit → duwit*                              *Javanese*
            Transpose (first, successor) over Onset

      c.    *balaynun → nulayban*                        *Hanunoo*
            Transpose (first, polar) over Body

      d.    *kenkänsä polk*i → *ponkansa kelki*          *Finnish*
            Transpose (first, doppler) over Body

Given these restrictions on ludlings to transposition operations and to total inversion, one would expect in ludling acquisition that the most important positions are the first and the last one. Transformations where items in these positions are switched may thus be more acceptable than transformations involving reference to absolute or relative position of non-edge syllables. This would explain why transformations (3i) and (3ii) are attested, while (3iii) and (3iv) are not. Moreover, if learners predominantly attend to the first and the last syllable, then even the choice between total reversal (3i) may not be much more acceptable than (3ii). We investigated these predictions empirically.

In the experiment, participants were first informed that they would witness a 'Martian rite'. In this rite, a chief Martian always pronounces a sentence, to which a subordinate Martian has to reply appropriately. Participants were also informed that these two Martians mastered the rite perfectly, and were instructed to try to figure out what the rite was about. Participants were presented with 25 trials, in which one synthesized voice (the chief Martian) pronounced a three syllable sequence and another synthesized voice (the subordinate Martian) replied with the same syllables but in reverse order.

After familiarization, participants were informed that they would witness the rite now with the chief Martian and another subordinate Martian who masters the rite less well. They were instructed to judge on a scale from 1 to 9 whether the new subordinate Martian's response conformed to the rules of the rite. They were instructed to press 1 if they were certain that the Martian's reply was wrong, 9 if they were certain that it was correct, and 5 if they were unsure. Then they completed 20 trials in which the chief Martian uttered a four-syllable sequence, and the new subordinate Martian replied with the same syllables in one of four different orders. In five of the trials, he replied with a 'natural' transformation. In five trials, this transformation was a complete inversion of the chief Martian's sequence; in other five trials only the first and the last syllable were switched, while the middle syllables remained in place (that is, the order was transformed from 1234 to 4231). In the other trials, the subordinate Martian replied with an 'unnatural' transformation. Half of these transformations were of the form '1234 → 1432', and the remaining transformations '1234 → 3412'. All syllables were consonant–vowel (CV) syllables synthesized with the Mbrola speech synthesizer (Dutoit *et al.* 1996).

As shown in Figure 1, the ratings for natural transformations ($M = 6.42$, $SD = 1.02$) were significantly higher than for unnatural ones ($M = 3.72$, $SD = 1.88$), $F(1,11) = 20.43$, $p = 0.0009$. While natural transformations were rated significantly above 5 (the neutral point), $t(11) = 4.83$, $p = 0.0005$, unnatural ones were rated significantly below, $t(11) = 2.37$, $p = 0.0371$.

The ratings (1234 → 4321: $M = 6.72$, $SD = 1.53$; 1234 → 4231: $M = 6.12$, $SD = 1.20$) did not differ significantly between the natural transformations, $F(1,11) = 1.25$, $p = 0.288$, ns; the ratings of the unnatural transformation (1234 → 1432: $M = 3.23$, $SD = 1.71$; 1234 → 3412: $M = 4.20$, $SD = 2.19$), in contrast, differed, $F(1,12) = 7.91$, $p = 0.017$.
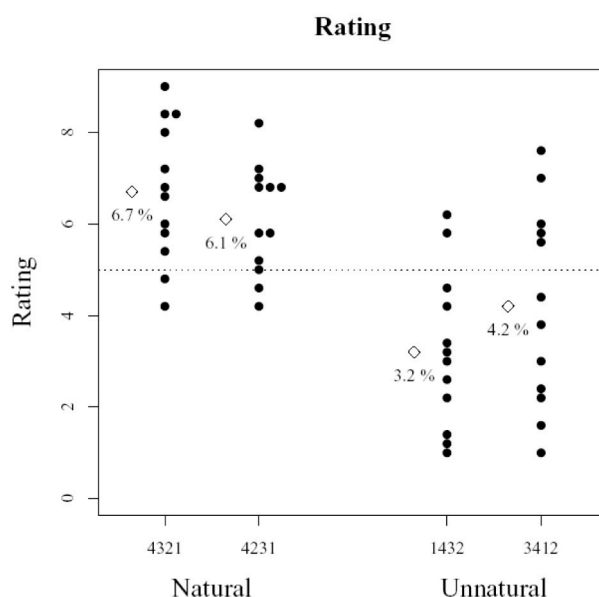
*Figure 1: Results of Nevins & Endress (2007; Exp. 1)*

These results clearly establish that the 'unnatural' hypotheses in (3iii) and (3iv) were not considered. There may have been a short-circuiting strategy that accounts for the numerical preference for (3iv) over (3iii), in that it is easier to detect that a transformation has not occurred when hearing σ1 in initial position.

The results are consistent with the hypothesis that natural transformations achieved by the operations in (5) are preferred to unnatural ones even though both types are logically 'consistent' with the data. They thus demonstrate an analytic bias in generalization over syllable-precedence transformations, one that exactly lines up with the typology of attested and non-attested extant ludlings.

One possible objection to our interpretation of these results is that they represent some kind of 'general sequence learning' and do not bear on the specific question of primitives of linguistic representation. To examine this possibility directly, we replicated the experiment with musical stimuli.

In a second experiment, the procedure was identical to that describe above, except that tones instead of syllables were used as stimuli. Before familiarization, participants were informed that they would witness a Martian rite, in which the chief Martian played a short melody, and a subordinate Martian had to reply appropriately with another melody. Then participants were familiarized with 30 trials in which the chief Martian played a four-tone melody on an instrument, and the subordinate Martian played its inversion on another instrument. The rationale for using four-tone melodies rather than three-item sequences as in Experiment 1 was that participants usually encode intervals among tones rather than their absolute pitches; in terms of intervals, however, we used again three-item sequences.

After this familiarization, participants were again informed that they would now witness the rite with the chief Martian, and another subordinate Martian who mastered the rules of the rite less well; they were instructed to rate the new Martian's performance on a scale from 1 to 9. The chief Martian (that is,

the same instrument as before) then played a five-tone melody comprising of 4 intervals (corresponding to the four-syllable sequences in Experiment 1). The new subordinate Martian then played a transformed melody in which the interval order (rather than the tone order) was transformed.

Moreover, since intervals are inverted when played backward (e.g., an upward octave becomes a downward octave), the intervals were also inverted. Again, the two natural transformations were 1234 → 4321 and 1234 → 4231, and the two unnatural transformations 1234 → 1432 and 1234 → 3412. Each transformation occurred five times in the test items.

As shown in Figure 2, participants rated the natural transformations ($M$ = 4.89, $SD$ = 1.24) better than the unnatural ones ($M$ = 4.17, $SD$ = 1.18), $F(1,12)$ = 11.96, $p$ = 0.006. However, participants rated the complete reversal ($M$ = 5.46, $SD$ = 1.47) better than the transformation 1234 → 4231 ($M$ = 4.32, $SD$ = 1.56), $F(1,12)$ = 5.70, $p$ = 0.034 and better than all other three as a group, $F(1,12)$ = 10.22, $p$ = 0.0077. Moreover, while the complete reversal was rated better than all other transformations (against 1234 → 4231: $t(12)$ = 2.39, $p$ = 0.0343; against 1234 → 1432: $t(12)$ = 4.05, $p$ = 0.0016; against 1234 → 3412: $t(12)$ = 2.33, $p$ = 0.0380), no other pair-wise differences were significant.
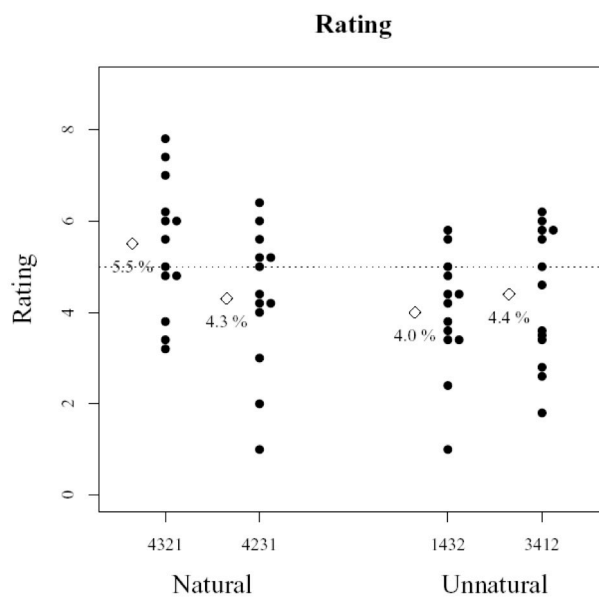


*Figure 2: Results of Nevins & Endress (2007; Exp. 2)*

When considered in light of the results of the experiment with linguistic stimuli, the results of the second experiment suggest that musical sequence transformations are not learned the same way as linguistic transformations. One possible explanation is that melodies (in particular atonal ones such as the melodies used here) may be encoded predominantly with respect to their *contours* (e.g., Dowling & Fujitani 1971); since all but transformation (3i) change the contour, one may expect that only transformation (3i) should be acceptable. Possibly, one may observe similar results using linguistic material that also features prosodic contours (e.g., suprasegmental tones). However, the question

may also be turned around to ask why edges are special in language but not music. While syllables bear intrinsic properties (such as their segmental content), musical notes largely function solely as links in a contour. While future research may reveal whether the analytic biases for edges in linguistic computation found in our Experiment 1 follow from more basic representational properties of sequence learning, the fact that they did not emerge in our Experiment 2 would suggest that it is words or syllables in particular that implicate a domain-specific learning bias.

Jointly considered, the experiments here allow one to conclude that (i) not every logically possible generalization is actually followed by humans when learning syllable-precedence-modifying ludlings, and (ii) the possibility of edge-switch as the generalization may be unique to linguistic computation. Final & antepenult switch (3iii) and every-other-switch (3iv) cannot be generated using the restrictive primitives in (5), are not found in the typology of existing ludlings, and were not generalized by our participants. The absence of (3iii) and (3iv) in existing ludlings turns out to be a principled rather than accidental gap. The study of universals is thus informed not only by what is shared among the languages of the world, but also by what is missing.

Taken in tandem these two conclusions implicate an analytic bias towards using only certain types of elements in the structural description of syllable-level generalizations — namely left edge, right edge, and ∀ (all syllables in the domain) — which coincides with the typology of existing natural ludlings. Not every way of generalizing a pattern is equally likely, which arguably is a relief for the learner in the face of representationally ambiguous data.

## 3. Universal Asymmetries between Consonants and Vowels

In this second case study we examine a universal dispreference for consonantal repetition as opposed to vowel repetition, focusing on the typological rarity of vowel dissimilation as opposed to widespread biases against consonantal identity as revealed in statistical analyses and experimental tasks (Berkley 2000, Walter 2007).

The source of this universal asymmetry is related to a more general difference between consonants and vowels. Typological, acquisition, and experimental studies point towards different functional roles for consonants and vowels. Maddieson's (2005) paper in the *World Atlas of Language Structures* reveals that of 564 languages surveyed, all have more consonants than vowels in their inventory. Nazzi *et al.* (2009) find that consonants are more important for vowels in word learning, by showing that when French- and English-learning 30-month old infants must neglect either a consonantal feature or a vocalic feature (e.g., match /pide/ with either /tide/ or /tüde/) that they chose to neglect the vocalic feature. Consonants and vowels are not even learned the same way, as consonants display categorical perception (Eimas *et al.* 1971) while vowels display perceptual magnet effects (Kuhl 1991).

Nespor *et al.* (2003) observe that no language is the inverse of Semitic, having vocalic lexical roots and consonantal glue, Peña *et al.* (2000) find that

consonants are easier for learning word-like 'frames', Surendran & Niyogi (2006) find that consonants have three times the functional load of vowels, and Cutler *et al.* (2000) find that, given a word like *kebra*, experimental participants find it easier to convert the word to *kobra* than to *zebra*. Owren & Cardillo (2006). find that consonants are more important for word identification and vowels are more important for talker identification.

All of these findings point to the conclusion that consonants bear the brunt of building lexical skeleta, and that vowels have a different functional role as grammatical, rhythmic, and sociolinguistic glue. As Nespor *et al.* (2003) point out, distinctiveness between consonants within a word tends to be maximized, whereas distinctiveness between vowels within a word tends to be reduced. Thus, an important asymmetry concerns the types of phonological processes found in each. Among vowels, harmony, a rule creating sub-segmental identity, is very common, while dissimilation is extremely rare (occurring only among low vowels, Suzuki 1998). Within consonants, on the other hand, dissimilation is extremely common, while consonant harmony, while existent (Hansson 2001), is rarely of the iterative type found in vowel harmony.

Much like the study above we can ask the question of whether these typological findings are simply due to sampling error or whether they reflect true universals. In particular we can ask whether consonantal repetition is dispreferred compared to vocalic repetition. In Nevins & Toro (2007), we investigated this question experimentally with 18 Italian subjects, none of whom had rules of obligatory consonant repetition or obligatory vowel repetition in their language. Thus any differences found between these two conditions should reflect true analytic biases.

The first pattern was a consonant repetition language with a rule of adjacent repetitions of consonants in CVCVCV words, where C1=C2. Vowels were always frames of the form CaCuCE or CiCeCo. These words were played to participants in a continuous 10 minute stream, thus of the form: [...*mamukEsisekosakakusE*...] (where E represents a lax vowel), with a 25 ms pause between each word. Participants were told that it was a broadcast of an alien language and that they should listen carefully. After 10 minutes of familiarization we tested participants in 16 trials on two forced-choice tests: One was a recognition task to see if participants recalled the vocalic frames. Thus subjects were asked whether *mamukE* or *mumeki* could be a word in the alien language that they heard. Participants thus only had to pay attention to the transitional probabilities among the stimuli in order to successfully discriminate between items that they recognized.

In the generalization test, participants were asked 'which of these could be a word in the language you heard: *Babure* or *ribero',* where both words respected the vocalic frames presented during exposure, but in one of which C1=C2 and in the other C1=C3, which was incongruent with the pattern presented during exposure. According to the distinctness-of-skeleta hypothesis, this rule should be hard to learn. Results of both tests are shown in Figure 3 below, with means indicated by a triangle.
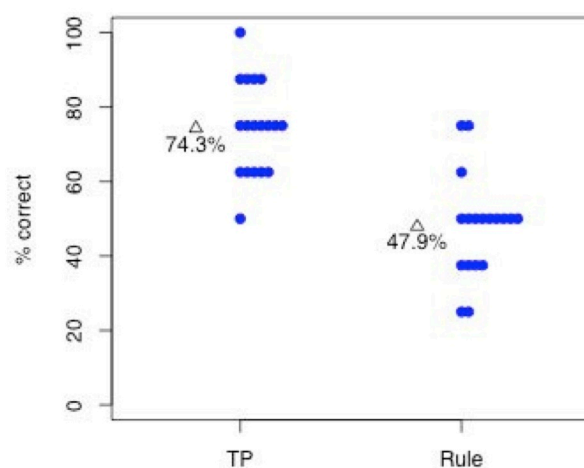
*Figure 3: Results of Nevins & Toro (2007; Exp 1)*

In the second experiment, we reversed the pattern. Consonants were the transitional-probability glue (mVkVfV or bVsVrV), and rules of adjacent repetition were defined over vowels where V1=V2 [...*mekefubisiromikifemobosa*...]. The procedure was otherwise identical to that of the first experiment. In a recognition task, participants were asked whether *mekefu* or *kefebu* was a possible word, and in the generalization task they were asked whether *makafu* with V1=V2 or *busaru* with V1=V3 was a possible word. Results are presented below, with means indicated by a triangle.
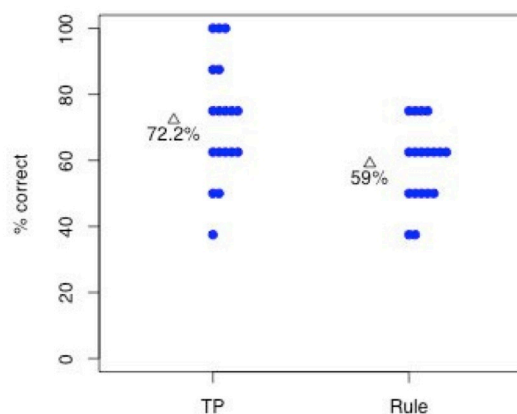


*Figure 4: Results of Nevins & Toro (2007; Exp 2)*

Recognition tasks did not differ from each other in the two conditions, $t(34) = -.40$, $p = 0.69$. However, the results of the two generalization tasks did, with generalization of the vowel repetition rule ($M= 59\%$, $SD = 11.9$) much better than generalization of the consonant repetition rule ($M=47\%$, $SD = 13.7$), $t(34) = 2.59$, $p = 0.014$. In the consonant-repetition condition, participants' discrimination of which of two stimuli 'belonged to the alien language' did not differ from chance, $t(17) = -0.64$, $p = 0.53$. By contrast, in the vowel-repetition condition, participants discrimination of which of two stimuli belonged to the alien language displayed a significant difference from chance, $t(17) = 3.20$, $p = 0.005$. In sum, people can

learn a repetition rule over vowels much better than over consonants.

Again, this experimental methodology with segmental processes of vowel and consonant anti-identity effects points to the same type of conclusion as in the syllabic processes of precedence-modification discussed above: The universal patterns observed through typological sampling of the world's languages can be tested in the laboratory to see if certain gaps are accidental or principled. Across the globe, there is a dispreference for consonantal repetition within words, and this same bias can be observed in failure to generalize during an artificial grammar experiment.

## 4. Concluding Remarks

The pursuit of universals must involve a three-fold approach: Rigorous typological sampling in order to catalogue what types of patterns are more common than others, formal modeling of the computational primitives that allow or favor one type of pattern over another, and experimental testing of whether the observed typological asymmetries and concomitant analytic biases are upheld in testing situations in which participants have no reason, other than Universal Grammar, to favor one type of pattern over another.

Returning to a parallel with the relation between possible and attested form in the study of organisms, Hauser (2009: 190–191) discusses the relevance of experiments such as those of Abzhanov (2004), based on beak shape in the Galapagos finches. In a certain sense, experimental genetic manipulations of organisms may be seen as analogous to the manipulations occurring in ludlings, insofar as the former are an attempt extend, modify, or exaggerate attested natural growth processes, much as the latter extend, modify, or exaggerate attested natural language processes. In these experiments, the genes encoding the proteins responsible for beak growth in large-beaked finches, bone morpho-genetic protein 4 (BMP4), were inserted into a chicken embryo, in order to understand the genetic primitives that lead to possible forms in nature. The result of these experiments is the smoothly unfolding development of chick with a large, broad beak, instead of the small beak that is typical of chickens. Hauser's (2009: 190–191) conclusion is that these studies underscore the importance of experimentation to understand constraints on organism form, as a complementary strategy to typological and naturalistic observation:

> It also shows why cataloguing variation in living animals is insufficient for understanding both the range of variation and its potential constraints; experimental studies such as those with chickens are necessary to uncover the limits of variation.

In this article we have discussed the importance of experimentation with invented ludlings for understanding two levels of phonological structure, intersyllabic processes and intersegmental processes, and attempted to demonstrate that two proposed universals — one derived from the seemingly obscure domain of ludlings and one derived from the well-known dualistic division between consonants and vowels — are both upheld in experimental

scenarios in which historico-cultural diachronic contingencies are rendered irrelevant, and in which the only remaining explanation for the observed linguistic asymmetries remains profoundly cognitive.

## References

Abzhanov, Arhat, Meredith Protas, B. Rosemary Grant, Peter R. Grant & Clifford J. Tabin. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305, 1462–1465.

Bagemihl, Bruce. 1989. The Crossing Constraint and 'backwards languages'. *Natural Language and Linguistic Theory* 7, 481–529.

Berkley, Deborah. 2000. Gradient obligatory contour principle effects. Chicago, IL: University of Northwestern PhD dissertation.

Boeckx, Cedric & Massimo Piattelli-Palmarini. 2005. Language as a natural object — Linguistics as a natural science. *The Linguistic Review* 22, 447–466.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Conrad, R. 1960. Serial order intrusions in immediate memory. *British Journal of Psychology* 51, 45–48.

Cutler, Anne, Nuria Sebastián-Gallés, Olga Soler-Vilageliu & Brit van Ooijen. 2000. Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition* 28, 746–755.

Dowling, William J. & Diane S. Fujitani. 1971. Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America* 49, 524–531.

Dutoit, Thierry, Vincent Pagel, Nicolas Pierret, François Bataille & Olivier van der Vreken. 1996. The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of the 4th International Conference on Spoken Language Processing*, vol. 3, 1393–1396.

Ebbinghaus, Hermann. 1885/1913. *Memory: A Contribution to Experimental Psychology*. New York: Teachers College, Columbia University. [http:// psychclassics.yorku.ca/Ebbinghaus]

Eimas, Peter D., Einar R. Siqueland, Peter W. Jusczyk & James Vigorito. 1971. Speech perception in infants. *Science* 171, 303–306.

Endress, Ansgar D., Brian J. Scholl & Jacques Mehler. 2005. The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, 134, 406–419.

Greenberg, Joseph H. 1978. *Universals of Human Language*. Stanford, CA: Stanford University Press.

Halle, Morris. 2008. Reduplication. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 325–357. Cambridge, MA: MIT Press.

Hansson, Gunnar Ólafur. 2001. Theoretical and typological issues in consonant harmony. Berkeley, CA: University of California dissertation.

Hauser, Marc D. 2009. The possibility of impossible cultures. *Nature* 460, 190–196.

Henson, Richard. 1998. Short-term memory for serial order: The start-end model.

*Cognitive Psychology* 36, 73–137.

Hicks, R. Yale, David T. Hakes & Robert Young. 1966. Generalization of serial position in rote serial learning. *Journal of Experimental Psychology* 71, 916–917.

Hitch, Graham J., Neil Burgess, John N. Towse & Vicki Culpin. 1996. Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 49, 116–139.

Kegl, Judy, Ann Senghas & Marie Coppola. 1999. Creation through contact: Sign language emergence and sign language change in Nicaragua. In Michel DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*, 179–237. Cambridge, MA: MIT Press.

Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105: 10681–10686.

Kuhl, Patricia K. 1991. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50, 93–107.

Laycock, Don. 1972. Towards a typology of ludlings, or play-languages. *Linguistic Communications* 6, 61–113.

Maddieson, Ian. 2005. Consonant–vowel ratio. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 18–21. Oxford: Oxford University Press.

Moro, Andrea. 2008. *The Boundaries of Babel*. Cambridge, MA: MIT Press.

Musso, Mariachristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel & Cornelius Weiller. 2003. Broca's area and the language instinct. *Nature Neuroscience* 6, 774–781.

Nazzi, Thierry, Caroline Floccia, Bérangère Moquet & Joseph Butler. 2009. Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology* 102, 522–537.

Nespor, Marina, Marcela Peña & Jacques Mehler. 2003. On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio* 2, 221–247.

Nevins, Andrew & Ansgar Endress. 2007. The edge of order: Analytic bias in ludlings. *Harvard Working Papers in Linguistics* 12, 43–53.

Nevins, Andrew & Juan Manuel Toro. 2007. Consonants as skeleta, vowels as glue. Paper presented at the *Ljubljana Linguistics Circle*. (University of Ljubljana, 21 May 2007)

Ng, Honey L.H. & Murray T. Maybery. 2002. Grouping in short-term verbal memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology: Section A* 55, 391–424.

Ohala, John J. 1986. Consumer's guide to evidence in phonology. *Phonology Yearbook* 3, 3–26.

Owren, Michael J. & Gina C. Cardillo. 2006. The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *Journal of the Acoustical Society of America* 119, 1727–1739.

Plénat, Marc. 1995. Une approche prosodique de la morphologie du verlan. *Lingua* 95, 97–129.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Piscataway, NJ: Rutgers University Center for Cognitive Science Technical Report 2.

Schulz, Rudolf W. 1955. Generalization of serial position in rote serial learning. *Journal of Experimental Psychology* 49, 267–272.

Smith, Neil V. & Ianthi Maria Tsimpli. 1995. *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.

Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In Ole Nedergaard Thomsen (ed.), *Competing Models of Language Change: Evolution and Beyond*, 43–58. Amsterdam: John Benjamins.

Suzuki, Keiichiro. 1998. A typological investigation of dissimilation. Arizona: University of Arizona doctoral dissertation.

Walter, Mary Ann. 2007. Repetition avoidance in human language. Cambridge, MA: MIT dissertation.

*Andrew Nevins*
*University College London*
*Division of Psychology and Language Sciences*
*Chandler House 115a*
*2 Wakefield Street*
*London, WC1N 1PF*
*United Kingdom*
*a.nevins@ucl.ac.uk*

# Syntax: Its Evolution and Its Representation in the Brain

Ljiljana Progovac

Poeppel (2008) observes that there is no clear correspondence between units of analysis in linguistics (especially the abstract and arbitrary-looking principles of syntax) and biological units of neuroscience, concluding that current neurolinguistic research presents a case of cross-sterilization, rather than cross-fertilization. Here the proposal is developed that decomposing syntax into intermediate evolutionary layers, into its evolutionary primitives, not only makes syntax compatible with gradualist accounts, but it also renders it more tangible and less abstract. In this approach, at least some complexities (and oddities) of syntax, such as Subjacency effects and the small clause core, can be seen as side-effects/by-products of evolutionary tinkering. It is conceivable that such evolutionary considerations are a necessary missing ingredient in any attempt to establish links between the postulates of syntax and the units of neuroscience. This article considers concrete linguistic data and suggestions as to where and how to look for neurobiological correlates of syntax.

*Keywords:*    evolution of syntax; ergativity; small clauses; subjacency

> *Nothing in biology makes sense except in the light of evolution.*
> (Dobzhansky 1973)

## 1.    Introduction

Poeppel (2008) observes that there is no clear correspondence between units of analysis in linguistics (especially syntax) and biological units of neuroscience, concluding that current neurolinguistic research presents a case of cross-sterilization, rather than cross-fertilization (see also Poeppel & Embick 2005).

Quite a bit is now known about the units of neuroscience: that neurons receive signals from other neurons through their dendrites, and transmit their own signals to other neurons through their axons; that signals are passed between neurons by synapses; that the human cortex includes around $10^{10}$ neurons, and that each of these neurons has $10^4$ synapses; and so on, and so forth. The theory of syntax has likewise led to significant accumulation of knowledge and crosslinguistic generalizations. But it is indeed hard to see how to match the units such as axons, neurons, or synapses, with the postulates of syntactic theory, such as Merge, Move, Theta Criterion, EPP, Subjacency. Nor is it likely that direct correlations of this kind will be found.

At the same time, language/syntax has to be represented in the brain somehow, and some new findings point in this direction. Pulvermüller (2002, this volume) argues that there is a neurobiological basis for words and sentences in terms of neurons. For him, language mechanisms are organized as nerve cells and their mutual connections. To cite one application, concrete words referring to objects and actions are proposed to be organized as widely distributed cell assemblies composed of neurons in sensory and motor areas. In contrast, highly abstract grammatical function words and affixes are assumed to be more focally represented in the left-hemispheric core language areas of Broca and Wernicke (Pulvermüller 2002: 49). His assumption is that there are one-to-one correspondences between linguistic representations and neuronal entities and between linguistic processes and neuronal processes (p. 209).

Ullman (2008) argues that specific language processes, such as irregular vs. regular past tense formation, activate two different types of memory: Declarative and procedural, respectively (see also Pinker & Ullman 2002). When it comes to language, declarative memory specializes for the storage of the Lexicon, including irregular morphology. Procedural memory, on the other hand, specializes for syntax and regular morphology, including sequences and rules, implicit knowledge, and rule-governed hierarchical (de)composition of complex forms. It is also of significance that the two memory systems overlap to some extent: Both can learn some of the same types of knowledge or skills, but with different computational and neural bases (see also Wray 2002: sect. 2.3).

The question is then not so much whether language/syntax is represented in the brain, but whether we can hope to find more direct correlates between the units of syntax and those of neuroscience. I suggest that one needs to explore a route which has, surprisingly, not been explored seriously — to challenge syntactic theory to decompose its postulates into more primitive entities, which would stand a better chance of being commensurate not only with the units of neuroscience, but also with the notions of evolutionary biology.[1]

Many properties of present-day syntax *look* arbitrary and abstract, including the two discussed in this article (Subjacency and the small clause core of clauses/sentences), leading to a wide-spread view among syntacticians that a gradualist evolutionary approach to syntax is impossible: Its principles are just too abstract for evolutionary forces to target them (e.g., Bickerton 1990, 1998,

---

[1]    In this respect, Kinsella (2009) calls for a syntactic theory which would be compatible with adaptive evolutionary processes.

Lightfoot 1991, Chomsky 2005). My contribution stands this argument on its head, and proposes that decomposing syntax into intermediate evolutionary layers, into its evolutionary primitives, not only makes syntax compatible with adaptationist accounts, but it also renders it more tangible and less abstract. In this approach, at least some complexities (and oddities) of syntax can be seen as side-effects/by-products of evolutionary tinkering.[2] It is conceivable that such evolutionary considerations are a necessary missing ingredient in any attempt to establish links between the postulates of syntax and the units of neuroscience.

Section 2 explores the nuts and bolts of the proto-syntax proposal. I first discuss the significance of the commonly accepted analysis according to which every modern clause/sentence unfolds from the small clause core, then explore the possibility that transitive constructions were tinkered out of the intransitive ones, leading to two types of present-day languages (ergative–absolutive and nominative–accusative), and finally consider the consequences of the proposal that proto-syntax was based on (intransitive) small clauses. Section 3 mentions some corroborating evidence for this view from language acquisition, aphasia, and genetics. Section 4 considers another linguistic universal, Subjacency, and offers a novel way of looking at it, consistent with the evolutionary proposal explored here. Section 5 concludes this article.

## 2.    Proto-Syntax and the Small Clause Universal

### 2.1.    Small Clause Core in the Light of Evolution

According to the mainstream syntactic theory, Minimalism and its predecessors, a clause/sentence is derived from a 'small clause' construct, an argument-predicate combination which typically excludes clausal functional projections.[3] This core subsequently unfolds/transforms into a full (finite) sentence/clause, after the Merge of, for example, Tense, and Move of the subject into the specifier of the tense phrase (TP).[4] The English sentences, or TPs, in (1) are thus derived from small clause structures such as (2), as illustrated in (3)–(5):

(1)    Maria will stay. / Maria was angry. / Maria is in Bamberg.

---

[2]    On recursion and epiphenomenal nature of certain syntactic principles, see also Arsenijević & Hinzen (this volume), Nevins (this volume), and Zeijlstra (2008).

[3]    The basic argument of this article focuses on intransitive structures — transitive predicates involve additional layers of structure and are probably later syntactic innovations. Section 2.2 offers some rationale for this view, as well as suggestions regarding how transitivity may have arisen, and what consequences this transition may have had on variation in present-day languages.

[4]    As pointed out by a reviewer, there are languages for which a TP-analysis may not be desirable, such as ergative languages or the so-called non-configurational languages. While resolving this issue is beyond the scope of this article, suffice it to say that the proposal here is that the small clause core is what all languages share. The way complexity is built above and beyond the small clause core may vary significantly across languages. For some speculation regarding the emergence of transitivity in ergative vs. nominative/accusative languages, see section 2.2; see also fn. 6 regarding some ancient languages.

(2) Maria stay. / Maria angry. / Maria in Bamberg.

(3) a. Small Clause: [$_{SC}$ Maria stay] →
    b. [$_{TP}$ will [$_{SC}$ Maria stay]]] →
    c. TP: [$_{TP}$ Maria [$_{T'}$ will [$_{SC}$ ~~Maria~~ stay]]]]

(4) a. Small Clause: [$_{SC}$ Maria angry] →
    b. [$_{TP}$ was [$_{SC}$ Maria angry]]] →
    c. TP: [$_{TP}$ Maria [$_{T'}$ was [$_{SC}$ ~~Maria~~ angry]]]]

(5) a. Small Clause: [$_{SC}$ Maria in Bamberg] →
    b. [$_{TP}$ is [$_{SC}$ Maria in Bamberg]]] →
    c. TP: [$_{TP}$ Maria [$_{T'}$ is [$_{SC}$ ~~Maria~~ in Bamberg]]]]

(The strikethrough notation in (3)–(5) and elsewhere indicates the original, pre-Move position of the subject.)

This idea has remained influential and widely accepted in syntactic theory ever since it was first proposed (e.g., Burzio 1981, Stowell 1981, 1983, Kitagawa 1986, Koopman & Sportiche 1991, Hale & Keyser 2002, Chomsky 1995, and subsequent minimalist work).[5] While languages and analyses vary with respect to what type, or how many, functional projections build on top of the small clause, most would agree that the small clause core is a universal property. Why is this so? Why should every sentence, in every language, be built upon the foundation of the small clause?

This can be explained by evolutionary tinkering, if all present-day human languages share a common evolutionary stage, that of small clause proto-syntax. In this view, the building of the modern clause e.g. in English involves (at least) two (semi-autonomous) systems/layers, as well as their complex interaction: An ancient system, and a more recent one.[6] In this scenario, TP/sentence would not

---

[5]  As far as I can tell, Pollard & Sag's (1994) criticism of Stowell's structures applies to those small clauses which are embedded within other clauses but does not necessarily extend to the small clauses from which a clause unfolds, or to the root small clauses, such as the ones illustrated in (17), and which obviously have to be constituents of some kind. For some discussion of the structure of embedded small clauses, and some thoughts on why and how they differ structurally from the small clauses discussed in this article, see section 2.3. Very roughly speaking, small clauses embedded within other clauses get integrated into the matrix clause by various morpho-syntactic processes, so that they show complexities not attested with root small clauses, and their constituency may be distorted by the movement of the small clause subject to (a functional projection inside) the main clause.

[6]  As pointed out by a reviewer, ancient languages may have relied more heavily on small clause syntax than modern languages do. In this respect, Latin is famous for its Absolute Ablative (*ablativus absolutus*) constructions, which basically involve small clauses tagged onto finite clauses; comparable constructions exist in modern English as well (see, e.g., Stump 1985 and also Progovac 2009b):

(i)  [Urbe capta] Aeneas fugit.                    *Latin*
     *city captured Aeneas fled*
     'With the city captured, Aeneas fled.'

have arisen from scratch, designed in an optimal way (e.g., Chomsky 2005), but rather it would have been tinkered from/superimposed upon what was already there: the small clause foundation, leading to quirks and complexities that syntax is (in)famous for (Progovac 2008a, 2009a). This approach is gradualist in nature, assuming progression in stages, and is thus in the spirit of Pinker & Bloom (1990) and Jackendoff (1999, 2002).[7]

Evolutionary explanations invoking layering and recency dominance can be found elsewhere, for example, in symbolic reference (Deacon 1997), in the superimposition of timed speech over ancient prosody (Deacon 1997, Pulvermüller 2002), in brain stratification accounts (in Vygotsky's and Piaget's work as well as in the triune brain proposals, such as MacLean 1949). The common theme in all is the inclusion of attainments of earlier stages in the *structures* of later stages, the theme which I explore here for the evolution of syntax.

## 2.2.  *Transitivity: An Excursus*

Transitive clauses involve additional layers of structure, and can be hypothesized to have been a later evolutionary innovation. While this article concentrates on intransitive predicates for this reason, the reviewers are correct in pointing out that the exclusion of transitivity needs more justification. This section is written in that spirit. It addresses the questions of how transitivity is treated in syntax and how it could have been tinkered over time from intransitive predicates, as well as how unaccusativity and ergativity may have developed given this evolutionary scenario.

In Minimalism (Chomsky 1995 and subsequent work), it is typically assumed that a transitive clause necessarily involves a $v$P/VP shell, that is, two verbal projections in which the arguments of the verb are generated, while intransitive structures, especially those involving theme arguments (unaccusatives), need not have the $v$P-layer:

---

In addition, on the basis of Vedic, Greek and Old Irish evidence, Kiparsky (1968: 51) has argued convincingly that proto-Indo-European, as well as early IE, was characterized by a frequent use of tenseless/moodless (injunctive) forms of the verbs, even in what we would consider today to be finite contexts (for the connections between these, on the one hand, and the small clauses and compounds, on the other, see Progovac 2006; also Progovac 2010b). In this respect, Gonda (1956: 36–37) notices that any attempt exactly to translate these injunctive categories into a modern Western idiom is doomed to fail, given "the vagueness in meaning and the great, and in the eyes of modern man astonishing, variety of its functions […]. [It] must sometimes be translated by a past tense […], sometimes by a present […] or future […], sometimes by a wish or command". Deutscher's (2000) discussion of Akkadian is also relevant in this respect. The issue certainly deserves further attention.

[7]   Pinker & Bloom (1990) assume the Baldwin Effect, the process whereby environmentally-induced responses set up selection pressures for such responses to become innate, triggering conventional Darwinian evolution. Tiny selective advantages are sufficient for evolutionary change: A variant that produces on average 1% more offspring than its alternative allele would increase in frequency from 0.1% to 99.9% of the population in just over 4,000 generations. This would still leave plenty of time for language to have evolved: 3.5–5 million years, if early Australopithecines were the first talkers, or, as an absolute minimum, several hundred thousand years, in the unlikely event that early Homo Sapiens was the first. Fixations of different genes can go in parallel.

(6)     Maria rolled the ball.

(7)     a.     [$_{\text{SC/VP}}$ rolled the ball]                                        →
        b.     [$_{vP}$ Maria [$_{\text{SC/VP}}$ rolled the ball]]                        →
        c.     [$_{\text{TP}}$ Maria [$_{vP}$ ~~Maria~~ [$_{\text{SC/VP}}$ rolled the ball]]]

If the transitivity layer is not there, the theme *the ball* will raise to become the subject of the TP:

(8)     The ball rolled.

(9)     a.     [$_{\text{SC/VP}}$ rolled the ball]]
        b.     [$_{\text{TP}}$ the ball [$_{\text{SC/VP}}$ rolled ~~the ball~~]]

Progovac & Locke (2009) propose that intransitive clauses antedated transitivity, based on the analysis of so-called exocentric verbal compounds, such as *daredevil, pickpocket, killjoy, rattlesnake, crybaby*. The make-up of these compounds, which they argue to be 'living fossils' of proto-syntax, lead them to conclude that proto-clauses involved verb-noun structures in which the noun's thematic role was syntactically underdetermined, and largely left open to pragmatic interpretation.

While the noun in these compounds is typically a theme (affected entity) (e.g., *pickpocket*), it could also be an agent or some other role (*crybaby*), or even vague with respect to these possibilities.[8] For example, Serbian *pali-drvce* ('ignite-stick, matches') is both a stick that ignites and a stick that gets ignited. Even though English *rattlesnake* happens to refer to a snake that rattles, it would not be impossible to imagine this label used to refer to people who rattle snakes, say as a hobby, on analogy with *scarecrow* (that which scares crows), *pickpocket* (one who picks pockets), etc.[9] Importantly, this dual possibility is not available with syntactically more complex compounds, such as *snake-rattler*, which involve an agentive suffix *–er* and with it a transitivity layer, and thus can only have the latter interpretation.

In addition to the verb-noun compounds discussed above, there are other constructions across languages which seem to still exhibit these simpler ('fossil') structures, in which the thematic role is left syntactically unspecified, and thus open to pragmatic interpretation. Consider, for example, the intransitives in ergative/absolutive languages, as illustrated in the following example from Tongan (Tchekhoff 1979: 409):

---

[8]     The proposal differs somewhat from that of Casielles & Progovac (2010), who also propose that intransitive (thetic) structures evolved prior to transitive structures. The difference is that in Casielles & Progovac the conclusion is that the expression of themehood (unaccusativity) preceded the expression of agenthood, the latter associated with the *v*P-projection. It may be that the morphosyntactic differentiation between the theme and agent, as attested in the data discussed in Casielles & Progovac, represented a somewhat later evolutionary development. The issue deserves further attention.

[9]     The so-called exocentric compounds of this type are attested in a variety of (unrelated) languages, with similar images, although they ceased to be productive in most (for details and examples, see Progovac & Locke 2009 and Progovac, to appear).

(10)  'oku kai  'ae   iká.                                                      *Tongan*
      *PRES eat   the   fish*
      'The fish eats.' / 'The fish is eaten.'

The syntax leaves it unspecified whether the only argument *the fish* of the intransitive verb above is the agent or the patient/theme (see also Gil 2005 for a discussion of comparable structures in Riau Indonesian). It is only through the addition of an agent argument (e.g., *the man*), presumably in the vP layer, that the role of *the fish* would be disambiguated to necessarily represent a patient/theme. The addition of the agent morpheme (*–er*) in verb-noun compounds has a comparable effect, as illustrated above for the compounds such as *rattlesnake* vs. *snake-rattler*.

What characterizes ergative languages, in contrast to nominative/ accusative languages, is that the subject of an intransitive predicate is morphosyntactically equivalent to the object of a transitive predicate (see, e.g., Dixon 1994). This kind of syntax also seems to characterize the exocentric compounds discussed above. Adopting the protosyntactic proposal advocated in Progovac & Locke (2009) and Progovac (to appear), one can envision the subsequent development of two types of languages, nominative–accusative and ergative–absolutive. With ergative–absolutive languages, the only argument in intransitive structures will remain marked (or unmarked) with the same absolutive case, regardless of its theta role, while the special marking (ergative) will be reserved for the argument introduced in the higher (innovative) structural layer — say, *v*P for concreteness. In nominative–accusative languages, on the other hand, the only argument in intransitive structures has to be associated with the same higher functional layer (TP) with which the highest argument of a transitive predicate is associated (see examples (6) and (8)), thus rendering the two indistinguishable on the surface.

As pointed out by a reviewer, a transition from intransitive to transitive structures is also clearly observed in the emergence of Nicaraguan Sign Language (NSL), which developed spontaneously by deaf children in the 1970s and 1980s (see, e.g., Kegl *et al.* 1999). According to the authors, the early pidgin stages of NSL do not use transitive [NP V NP] constructions, such as (11) (Kegl *et al.* 1999: 216–217). Instead, the structure is typically broken into two (intransitive) clauses, [NP V NP V] sequences, with each verb taking only one argument, as illustrated in (12)–(14):

(11)  *WOMAN PUSH MAN.

(12)  WOMAN PUSH MAN GET-PUSHED.[10]

---

[10]    As pointed out by Kegl *et al.* (1999: 217), even though the gloss is passive in the second NP–V sequence, the form is not passive, but is rather marked with the first person point of view. As opposed to PUSH in the first NP–V sequence, where PUSH is articulated from the perspective of the pusher, the GET-PUSHED part is signed with "the signer's body jolting backwards, as if having received the thrust of a push by some unspecified agent". The GET-PUSHED type of data may be comparable to the examples such as *Problem solved*, *Case closed*, etc., which also appear to be passive, but may not be syntactically so, as discussed in fn. 15.

(13)   WOMAN PUSH MAN REACT.

(14)   WOMAN PUSH MAN FALL.

Typically, the primary meaning is expressed with the first verb, while the second verb tends to express the result or termination of the event. They found that second generation signers drop the second verb, creating constructions that can be seen as transitive [NP V NP] structures.

In fact, if transitivity indeed arose in a comparable fashion in the evolution of human language, then the *v*P/VP shell of modern syntactic theory can be seen as a ghost of this evolutionary past, which entertained structures with two verbs.

## 2.3.   *'Fossils' of Proto-Syntax*

There are good arguments for the small clause core analysis outlined in section 2.1. First, a full clause/sentence (TP) appears to have two subject positions, both of which can sometimes be overtly filled. In (15), there is an expletive (meaningless) subject in TP (*there*), and another subject (*a spider)* in the small clause, which agrees with the verb. In (16), one piece of the subject, *the spiders*, occurs in TP, while the remaining piece, the so-called floating quantifier (*all*), occurs in the small clause (see, e.g., Koopman & Sportiche, 1991):[11, 12]

(15)   [$_{TP}$ <u>there</u> was [$_{SC}$ <u>a spider</u> in the room]]

(16)   [$_{TP}$ <u>the spiders</u> were [$_{SC}$ <u>all</u> in the room]]

These data implicate two (partly overlapping) layers of clausal structure, each with a subject position (see Progovac 2008a, 2008b for more examples of such overlap in other languages, involving aspect, tense, and agreement). In an evolutionary framework, overlap and redundancy are unsurprising — as put in Carroll (2005: 170–171), "multifunctionality and redundancy create the opportunity for the evolution of specialization through the division of labor".

But perhaps most intriguing evidence for small clause constructs, typically not discussed in syntactic literature, comes from the use of such constructs in root contexts (see, e.g., Akmajian 1984, Roeper 1999, Potts & Roeper 2006, Progovac 2006, 2009a):

---

[11]   Comparable data can be found across languages. In (i), from Arabic, one conjunct is in TP (*Kareem*), while the rest of the conjunction is in the small clause (Aoun *et al*. 1994):

   (i)   Kariim   keen   huwwe   w Marwaan   çam   yilçabo.           *Arabic*
         *Kareem   was   he   and Marwaan   ASP   playing*
         'Kareem and Marwaan were playing.'

[12]   One should point out that the argument due to quantifier float in (16) is not conclusive given that there are alternative analyses of quantifier float which treat these quantifiers as adverbs (e.g., Kayne 1975, Bobaljik 1995). Bošković (2004) is a recent defense of the original proposal of quantifier float (16), which also attempts an explanation of the ungrammaticality of examples such as (i), brought up by a reviewer:

   (i)   * They arrived all.

(17)   Maria stay?! / Maria angry?! / Maria in Bamberg. (e.g., picture caption) /
       Him stay?! / Me late?! / Me first! / Family first! / Problem solved. /
       Case closed.

These small clauses, with arguably a single layer of clausal structure, can be seen
as 'living fossils' of a proto-syntactic stage in the evolution of human language,
with TP representing a later addition/innovation (Progovac 2008a, 2009a).[13] Put
another way, in the evolution of human language, clauses such as *Me late*
antedated clauses such a *I am late. / I will be late. / I might be late.* The existence of
the quirky clauses in (17), and the universal unfolding of clausal structure from
the underlying small clause (section 2.1), both begin to make sense if seen as
vestiges of gradual evolution of syntax, but remain mysterious otherwise.
       Uriagereka (2008) looks at embedded small clauses, such as the bracketed
clause in (18), and concludes that the structure of these (embedded) small clauses
is rather basic, and may involve finite-state syntax, the simplest type of syntax in
Chomsky's hierarchy. [14]

(18)   I cannot imagine [$_{SC}$ Maria angry].

One of the arguments Uriagereka invokes for the primitive nature of (embedded)
small clauses is the long-noted observation that these clauses do not have an
internal source of structural case for their subjects, which are thus assigned case
by an external element, the verb *imagine* in (18). Progovac (2006) argues that root
small clauses, of the type illustrated in (17), likewise do not have a structural

---

[13]   This is in the spirit of Jackendoff (1999, 2002) and Bickerton (1990, 1998), who claim that
       previous stages of evolution left traces/fossils in present-day languages, or continued to live
       in parallel with more complex structures ('living fossils'). According to Ridley (1993: 525),
       living fossils are species that have changed little from their fossil ancestors in the distant
       past, such as lungfish, for example.

[14]   A reviewer points out that there is a possibility for recursion in embedded small clauses,
       casting doubt on the view that such small clauses involve finite-state syntax:

          (i)      I consider [considering syntax boring] a mistake.

       However, recursion in small clauses seems more restricted (and forced) than recursion with
       e.g., finite clauses. In Progovac (2010a), I point out that there is a clear contrast between
       recursion in finite CP clauses (ii), which seems free and unlimited, and recursion in
       embedded small clauses (iii), which seems restricted to one or two levels of embedding:

          (ii)     Mary believes [that John knows [that the neighbors noticed [that he fell off his
                   motorcycle]]].

          (iii)  ?? I will let [John imagine [Peter see [Mike fall off his motorcycle]]].

       The argument there is that there is a continuity of clause complexity (from small clause to
       finite clause), which correlates with the continuity in recursion potential. Significantly for
       the arguments made in this article, small clauses which are not integrated into finite clauses,
       and which do not involve even structural case checking (see the discussion later in the text),
       do not allow recursion or embedding at all (Progovac 2010a):

          (iv)   a.   *Him worry [case closed].
                 b.   *Him worry [her happy [problem solved.]]

mechanism for checking case on their subjects, providing another argument that they are creations akin to embedded small clauses. Since with root small clauses there is no external source of case either, their subjects surface with what can be analyzed as default case, in the sense of for example, Schütze (2001) — witness the accusative on the pronominal subjects in (17). The evolutionary perspective explored here sheds light on the existence of both embedded and root small clauses, the latter typically not recognized as objects worthy of syntactic inquiry.

As pointed out by a reviewer, the embedded small clause in (19) seems to involve a displacement (Move) of the unaccusative subject to a position in which it checks accusative case (20), as has been more recently assumed in Minimalist literature (e.g., Bošković 2004; see Pollard & Sag (1994) and references there for equivalent and much earlier subject-to-object raising proposals in alternative frameworks).

(19)   I watched three men arrive.

(20)   a.   [$_{SC}$ arrive three men]          →
       b.   [$_{XP}$ three men [$_{SC}$ arrive ~~three men~~]]

The position into which the SC subject moves is often considered to be the same position in which the matrix object would surface, such as for example, Agr$_O$P position (Bošković 2004). On this analysis, the subject is not moving within the small clause, but rather out of the small clause and into a matrix clause position. Given this analysis, one can still maintain that the SC itself is a rather basic creation, with no functional layers of its own.

As argued in Progovac (2006), the main difference between embedded small clauses and root small clauses boils down to the following: While the former are integrated into the rest of syntactic structure by for example, Move and/or structural case checking involving the subject of the small clause, root small clauses such as the ones in (17) arguably involve no structural case on the subject, no Move, and only one layer of structure. This distinction correlates with the contrast below:

(21)   a.   Problem solved.
       b.   I want the problem solved.

The article is only required in the embedded small clause (21b), but not in the root small clause (21a), which correlates with the postulated structural case checking involving the subject of the small clause: While structural case requires the presence of a DP (Longobardi 1994), default case can be associated with mere NPs (see e.g., Schütze 2001 and examples such as *the real me*; for details of the analysis of root small clauses, see Progovac 2006). [15]

---

[15]   While I assume here and elsewhere that passive-like examples such as *Problem solved* and *Point taken* involve a single Merge and no Move, a reviewer is right to point out that this assumption needs defending, especially in the light of examples such as (i) below, provided by the reviewer:

Serbian unaccusative clauses provide an unambiguous argument for the empirical reality of root small clause syntax. Unaccusative verbs, i.e. intransitive verbs whose only argument is a theme (e.g., *arrive, fall, come, appear*), are analyzed cross-linguistically as starting/merging their subjects as complements/objects of the small clause (e.g., Perlmutter 1978, Burzio 1981):

(22)  a.  [$_{SC}$ arrived three men]]                                    →
      b.  [$_{TP}$ have [$_{SC}$ arrived three men]]]                      →
      c.  [$_{TP}$ three men [$_{T'}$ have [$_{SC}$ arrived ~~three men~~]]]

The unaccusative Serbian sentences in (23) would be derived from the corresponding unaccusative small clauses in (24), as illustrated in (25):

(23)  a.  Zima          je        stigla.                   *Serbian*
          *winter.FEM.3SG  AUX.3SG  arrived.FEM.SG*
          'Winter has come.'
      b.  Vlada              je          pala.
          *government.FEM.SG    AUX.3SG    fallen.FEM.SG*
          'The government has collapsed.'

(24)  a.  Stigla zima.
      b.  Pala  vlada.

(25)  a.  Small clause:  [$_{SC}$ stigla zima]]              →
      b.                 [$_{TP}$ je [$_{SC}$ stigla zima]]]  →
      c.  TP:   [$_{TP}$ zima [$_{T'}$ je [$_{SC}$ stigla ~~zima~~]]]

The prediction is, if there are small clause counterparts to these unaccusative clauses, then they should surface in the unaccusative VS word order. This is indeed the case in Serbian, as illustrated in (24).[16] Unaccusative small clauses are

---

(i)      Problem presumably solved.

The question here is where the adverb is attached. It is typically assumed in Minimalism (e.g., Cinque 1999) that adverbs are attached above the $vP$/VP shell and that arguments get to precede such adverbs only if they move to higher clausal projections. It is interesting in this respect that, unlike adverbs, parentheticals are not licit inside these small clauses (ii), in sharp contrast to the finite counterparts (iii):

(ii)     * Problem, I believe, solved.

(iii)    The problem, I believe, has been solved.

While (i) seems to call for a functional projection inside the small clause, (ii)–(iii) seem to argue against the presence of such a projection. The issue warrants further attention.

[16]  The closest counterpart in English would be the semi-fossilized unaccusative clauses, which necessarily surface in VS order, such as the underlined expression in (i):

(i)      <u>Come November</u>, he will go hunting.

According to a reviewer, *come* in (i) can be analyzed as a preposition recently grammaticalized from a verb. Even if so, the construction it grammaticalized from would have involved an unaccusative verb followed by its only argument.

more than just a theoretical construct in Serbian: They are in productive use, wearing the unaccusative syntax on their sleeve (for details see Progovac 2008b; see also Casielles & Progovac 2010 for comparable data from Spanish and other languages).

That (unaccusative) small clause syntax can have a life of its own is further confirmed by the existence of fossilized/formulaic clauses, which cannot even be expanded into full sentences (thanks to Ana Progovac, p.c., for bringing these to my attention). If expanded, they automatically acquire the literal, non-formulaic reading:[17]

(26)  a.    Pala  karta.                                                          *Serbian*
            *fallen  card*
            'Card laid, card played.'[18]
      b.    #Karta je pala.
            'The card fell.'

(27)  a.    Proš'o   voz.
            *gone      train*
            'The opportunity has passed.'
      b.    #Voz je prošao.
            'The train is gone.'

It is of relevance here that formulaic speech in general has been argued to be processed by the more ancient structures of the brain, showing resilience in cases of aphasia and other disorders (e.g., Code 2005 and Wray 2002).

Given this, processing of formulaic speech in the form of small clauses may provide a promising track to explore in neuroscience, one that can shed light on the distinction between what I postulate here to be (fossils of) proto-syntax and the more complex and more recent TP syntax. The production/perception of a TP may have to tap into two distinct neural mechanisms, with possibly some overlap: the one that supports the proto(-syntax) of small clauses, and another that supports the more recent TP syntax, necessarily activating the procedural memory. In other words, one may find neurobiological correlates of finiteness (TP expression) by comparing and contrasting the processing of small clauses (*Problem solved; Stigla zima*) with the processing of full finite clauses, such as *The problem has been solved; Zima je stigla*.) In addition, in the light of the discussion of transitivity in section 2.2., one may also expect to find neural correlates of transitivity by comparing and contrasting the processing of compounds such as *rattlesnake* with the compounds such as *snake-rattler*.

Furthermore, one may find that the proto-syntactic constructs, at least the formulaic ones, are more likely to be accessible to both procedural and declarative memories (see Pinker & Ullman 2002), as formulaic speech typically is (Code 2005, Wray 2002). Such constructs straddle the boundary between the lexicon and syntax, constituting good candidates for providing a transition from

---

[17]    The verbs in Serbian examples are past participles, which, unlike English translations might suggest, cannot be analyzed as adjectives.
[18]    Thanks to Ann Sawyer (p.c.) for the idiomatic translation.

a lexical stage (without syntax) to stages with more elaborate syntax.[19]

## 3.      Converging Corroborating Evidence

There is converging evidence from various disciplines suggesting that a simpler (non-TP) syntax underlies, and provides foundation for, our mastery of more complex syntax (see also fn. 6). Many have argued that language acquisition proceeds from a small clause stage to a TP stage (e.g., Radford 1988, Lebeaux 1989, Ouhalla 1991, Platzack 1990, Roeper 1999, Potts & Roeper 2006; but see Guasti 2002 for opposing views).[20] Kolk (2006 and references cited there) has argued that sub-sentential speech, including small clauses, requires less processing time (is processed within a smaller temporal window), and that it is thus frequently resorted to in agrammatic production as preventive adaptation.

Specific language impairment (SLI) is characterized, among other symptoms, by the delay or deficit in the use of auxiliary verbs, tense, and agreement morphology as well as of other functional categories, all potentially symptomatic of the lack of the TP layer. This kind of grammar thus resembles small clause grammar in relevant respects. Recently, a gene has been identified, FoxP2, whose mutation seems responsible for the disorder (Lai *et al*. 2001). According to Piattelli-Palmarini & Uriagereka (2005: 38), hominids in possession of a version of FOXP2 prior to the last mutation may have had a linguistic performance not unlike that of affected individuals. Given the approach explored in this article, one can hypothesize that these hominids would have been using a kind of proto-syntax characterized by the small clause constructs comparable to the ones illustrated in (17), (24), (26), (27), and (30).

## 4.      Another (Related) Universal: Subjacency

Subjacency is another principle central to syntax, taken to prohibit Move(ment) out of various 'islands,' including adjuncts and conjuncts, on which I focus here (see e.g., Ross 1967, Huang 1982, and Chomsky 1986):

(28)  *Who did Peter resign [after Mary met ~~who~~?]                    *adjuncts*

(29)  *Who did he hurt ~~who~~ and Mary found out?                       *conjuncts*

The current view of Subjacency in Minimalism and its predecessors is that Move is the default option, while Subjacency, restrictions on Move, is marked and in need of characterizing (Stepanov 2007, Chomsky 2008). This view feeds the influential language evolution hypothesis, according to which Merge (which

---

[19]    As pointed out by a reviewer, Construction Grammar frameworks advocate a continuum between lexicon and syntax (see e.g., Goldberg 1995: 7 and references there).

[20]    For some old and some recent views on the relationship between ontogeny/DEVO (development in children) and phylogeny/EVO (development in species), see for example, Studdert-Kennedy (1991), Ridley (1993), Carroll (2005), and Locke & Bogin (2006).

subsumes Move) was the only evolutionary breakthrough for syntax (e.g., Hauser *et al*. 2002, Chomsky 2005). Berwick's (1998: 338–339) words echo the common sentiment among syntacticians, that "there is no possibility of an 'intermediate' *syntax* between a non-combinatorial one and full natural language — one either has Merge in all its generative glory, or one has no combinatorial syntax at all".

But there is an alternative possibility (also mentioned in Cinque 1978, Bouchard 1984, Postal 1997, Boeckx & Grohmann 2007, and Progovac 2009b), that No Move is the default, and motivating Move a special, marked option. The constructions that prohibit Move are much more numerous and diverse than those that allow it, and they also do not form a natural class. If Subjacency is an elsewhere condition, a by-product of the evolution of syntactic complexity, rather than a principle of grammar, then it is unsurprising that there is still no satisfactory account of Subjacency (for discussion along these lines and references, see Progovac 2009b).[21]

My argument is that proto-syntax, based on small clauses introduced above, did not have Move or subordination/recursion (Progovac 2009b, 2010a; see also fn. 14). Initial clausal combinations arguably looked like paratactic constructs in (30). Indeed, neither root small clauses (31), nor their paratactic combinations (32), allow any manipulation by Move:[22]

(30)  a.    Nothing ventured, nothing gained.
      b.    Easy come, easy go.
      c.    Monkey see, monkey do.
      d.    No money, no come.

(31)  a.    *When problem solved?
      b.    *Whom worry?!

(32)  a.    *What ventured, nothing gained?
      b.    *Who monkey see, do?

---

[21]    As pointed out by a reviewer, there are also pragmatic approaches to Subjacency, such as Kuno (1987), for example. While pragmatics probably plays a role, it is also inevitable to conclude that syntactic structure plays an important role as well, especially given contrasts such as the one illustrated below, which would be difficult to reduce to pragmatics, but which clearly involve two different syntactic structures, one of which, coordination, as I argue, is more ancient and thus not subject to Move:

(i)    What did you eat the chicken with ~~what~~?

(ii)    * What did you eat the chicken and ~~what~~?

[22]    A reviewer points out that movement out of a small clause such as *When problem solved* does not occur because it seems impossible to have adverbials in such clauses in the first place (but see fn. 15 for seemingly possible use of adverbs in root small clauses). This still leaves us with evidence from the rest of the examples in (31)–(32), given that arguments and adverbials are equally affected. And, if indeed adverbs are not welcome in such clauses, this reinforces the view that these clauses are very rudimentary creations, arguably creations without functional categories.

According to, for example, Traugott & Heine (1991) and Deutscher (2000), grammaticalization of subordination proceeds through three stages: parataxis/ adjunction, coordination, and subordination, the older stages clearly being preserved alongside the innovations. If comparable stages/processes characterized language evolution (see e.g., Jackendoff 1999, Progovac 2009b, 2010a), then such evolutionary tinkering left us with multiple possibilities which partly overlap in function (see (33)–(35) below).

In this proposal, adjuncts and conjuncts are seen as older syntactic structures, less integrated into sentential fabric, which found their niche and continued to be used in parallel with subordination, a more recent innovation.[23] In this view, then, the reason why Move is prohibited out of adjunct or conjunct clauses is not because it would involve crossing some (combination) of structural barriers/boundaries, but rather because these are fossil structures, patched onto more complex syntactic structures, but still preserving their opacity with respect to Move. This would render these Subjacency effects epiphenomena of evolutionary tinkering. But does subordination bring any tangible novel possibility not afforded by adjunction or coordination?

Importantly, in addition to allowing Move, subordination also provides a recursive mechanism for embedding multiple viewpoints one within another, typically unavailable with either coordination or adjunction, privileging (35) over (33)–(34) in this respect:[24]

(33)   [As you know,] [as Mary knows,] he is a linguist.        *adjunction/parataxis*

(34)   He is a linguist, [and you know it,] [and Mary knows it]. *coordination*

(35)   You know [that Mary knows [that he is a linguist]].        *subordination*

If so, then subordination (and with it the possibility to apply Move across clauses) would have significantly increased the expressive power of language, in a concrete and tangible manner, constituting a plausible target for natural/sexual selection (see Progovac 2009b for a full(er) treatment of Subjacency along these lines).

---

[23]   Clausal conjuncts and adjuncts have been repeatedly noted *not* to be fully integrated into syntactic fabric, resembling separate utterances. First, they are often parsed as separate intonation-phrases (Nespor & Vogel 1986, Stowell 1981, An 2007). Next, adjuncts have been analyzed as merging in a different plane (Chomsky 2001), and conjuncts as sitting on parallel planes (Goodall 1987).

[24]   As brought up by a reviewer, Reis (1995: 53) argues that certain parenthetical constructions in German do allow the stacking of points of view, as in:

    (i)    Was     glaubst   du schätzt   er, wieviel   das   Auto   kosten wird?
           *what    believe    you estimates  he how.much   the    car    cost    will*
           'How much do you think he estimates that the car will cost?'

    Indeed, some of the German data discussed in the paper seem to be in transition, neither clearly integrated nor clearly parenthetical. While my claim is *not* that parentheticals can never express the stacking of points of view, it does seem that this kind of stacking becomes automatic, unambivalent, and streamlined only under subordination.

## 5.    Discussion and Conclusions

The central postulates of present-day syntax *look* arbitrary and abstract, including the two universals discussed in this article: the small clause core of every clause/sentence, and the islandhood of for example, conjuncts and adjuncts. This leads many syntacticians to conclude that a gradualist/adaptationist approach to syntax is impossible: The principles of syntax are just too abstract for evolutionary forces to target them. Similar considerations have led to Poeppel & Embick's (2005) conclusion regarding cross-sterilization between syntax and neuroscience (section 1). On the other hand, my proposal is that decomposing syntax into intermediate (evolutionary) layers not only makes syntax compatible with gradualist/adaptationist accounts, but it also renders it potentially commensurate with the units of neuroscience.

If proto-syntactic small clause constructs discussed in this article involve a simple concatenation grammar, perhaps just one instance of (Proto-)Merge (Progovac 2009a), then determining how they are processed in the brain, in contrast to their finite counterparts, can lead to important insights in neurolinguistics.[25] Indeed, separating out the relevant (evolutionary) layers may be necessary in formulating precise hypotheses regarding how syntax gets represented in the brain. The production/perception of a TP may have to tap into two distinct neural mechanisms, with possibly some overlap: The one that supports the proto(-syntax) of small clauses, and another that supports the more recent TP syntax, necessarily activating the procedural memory. In other words, one may find neurobiological correlates of finiteness (TP expression) by comparing and contrasting the processing of small clauses (*Problem solved; Stigla zima*) with the processing of full finite clauses, such as *The problem has been solved; Zima je stigla.*) In addition, one may expect to find neural correlates of transitivity by comparing and contrasting the processing of compounds such as *rattlesnake* with the compounds such as *snake-rattler*.

It is also of note that many root small clauses discussed in this article are formulaic expressions, the observation relevant not only for declarative and procedural memory considerations, but also for the finding that formulaic speech is processed by the more ancient structures of the brain. Thus, one may find that the proto-syntactic constructs, at least the formulaic ones, are accessible to both procedural and declarative memories (see Pinker & Ullman 2002), as formulaic speech typically is (Code 2005, Wray 2002).

---

[25]    As pointed out by a reviewer, the claim that proto-syntactic creations involve a basic concatenation of two elements finds support in various functionalist approaches to language. Many philosophers have emphasized the foundational nature of the subject/predication formations (e.g., Strawson 1964; see also Potts & Roeper 2006). More recently, Krifka (2008) has proposed that bimanual tool making might have constituted a pre-adaptation for topic/comment structures, noticing that the non-dominant hand can be likened to the topic of a sentence, and the dominant hand to the comment. Hurford (2007) and Casielles & Progovac (2010) discuss the significance of the topic-comment structures in the evolution of human language. Given the discussion in Casielles & Progovac (2010), it may be that topic-comment structures were preceded by comment-only structures (or wide-focus structures), often characterizing thetic unaccusative statements. The issue deserves further attention.

In order for the neurolinguistic research to produce cross-fertilization, rather than cross-sterilization, syntactic theory will have to turn aggressively to these important interfaces—syntactic representation in the brain, and evolution. The approach explored in this article is in that spirit. Even at first sketch, it reveals some new directions that may indeed pave the way to new discoveries.

## References

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *Natural Language and Linguistic Theory* 2, 1–23.

An, Duk-Ho. 2007. Clauses in non-canonical positions at the syntax-phonology interface. *Syntax* 10, 38–79.

Aoun, Joseph, Elabbas Benmamoun & Dominique Sportiche. 1994. Agreement, word order, and conjunction in some varieties of Arabic. *Linguistic Inquiry* 25, 195–220.

Arsenijević, Boban & Wolfram Hinzen. This volume. Recursion as a human universal and as a primitive.

Berwick, Robert C. 1998. Language evolution and the minimalist program: The origins of syntax. In Hurford *et al*. (eds.), 320–340.

Bickerton, Derek. 1990. *Language and Species*. Chicago, IL: University of Chicago Press.

Bickerton, Derek. 1998. Catastrophic evolution: The case for a single step from protolanguage to full human language. In Hurford *et al*. (eds.), 341–358.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Cambridge, MA: MIT dissertation.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. Putting phases in perspective. *Syntax* 10, 204–222.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language and Linguistic Theory* 22, 681–742.

Bouchard, Denis. 1984. *On the Content of Empty Categories*. Dordrecht: Foris.

Burzio, Luigi. 1981. Intransitive verbs and Italian auxiliaries. Cambridge, MA: MIT dissertation.

Carroll, Sean B. 2005. *Endless Forms Most Beautiful: The New Science of Evo–Devo and the Making of the Animal Kingdom*. New York: W.W. Norton and Co.

Casielles, Eugenia & Ljiljana Progovac. 2010. On protolinguistic 'fossils': Subject-verb vs. verb-subject structures. In Andrew D. M. Smith, Marieke Schouwstra, Bart de Boer & Kenny Smith (eds.), *The Evolution of Language: Proceedings of the 8th International Conference (EVOLANG 8)*, 66–73. Hackensack, NJ: World Scientific.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos P. Otero & Maria-Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1978. Towards a unified treatment of island constraints. *Innsbrucker Beiträge zur Sprachwissenschaft: Proceedings of the 12th International Congress of Linguistics*, 344–348.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Code, Chris. 2005. First in, last out? The evolution of aphasic lexical speech automatisms to agrammatism and the evolution of human communication. *Interaction Studies* 6, 311–334.

Deacon, Terrence W. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York & London: W.W. Norton & Company.

Deutscher, Guy. 2000. *Syntactic Change in Akkadian: The Evolution of Sentential Complementation*. Oxford: Oxford University Press.

Dixon, Robert M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dobzhansky, Theodosius. 1973. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* 35, 125–129.

Gil, David. 2005. Isolating-monocategorial-associational language. In Henri Cohen & Claire Lefebvre (eds.), *Handbook of Categorizaton in Cognitive Science*, 347–379. Amsterdam: Elsevier.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.

Gonda, Jan. 1956. *The Character of the Indo-European Moods (with Special Regard to Greek and Sanskrit)*. Wiesbaden: Harrassowitz.

Goodall, Grant. 1987. *Parallel Structures in Syntax: Coordination, Causatives and Restructuring*. Cambridge: Cambridge University Press.

Guasti, Maria Teresa. 2002. *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press.

Hale, Ken & Samuel Jay Keyser. 2002. *Prolegomena to a Theory of Argument Structure*. Cambridge, MA: MIT Press.

Hauser, Mark, Noam Chomsky & W. Tecumseh Fitch. 2002. The language faculty: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Huang, James. 1982. Logical relations in Chinese and the theory of grammar. Cambridge, MA: MIT dissertation.

Hurford, James R. 2007. *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Hurford, James R., Michael Studdert-Kennedy & Chris Knight (eds.). 1998. *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press.

Jackendoff, Ray. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3, 272–279.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Kayne, Richard. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kegl, Judy, Ann Senghas & Marie Coppola. 1999. Creation through contact: Sign language emergence and sign language change in Nicaragua. In Michel DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*, 179–237. Cambridge, MA: The MIT Press.

Kinsella, Anna R. 2006. *Language Evolution and Syntactic Theory*. Cambridge: Cambridge University Press.

Kiparsky, Paul. 1968. Tense and mood in Indo-European syntax. *Foundations of Language* 4, 30–57.

Kitagawa, Yoshihisa. 1986. Subjects in English and Japanese. Amherst, MA: University of Massachusetts dissertation.

Kolk, Herman. 2006. How language adapts to the brain: An analysis of agrammatic aphasia. In Progovac *et al.* (eds.), 229–258.

Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85, 211–258.

Krifka, Manfed. 2008. Functional similarities between bimanual coordination and topic/comment structure. In Regine Eckardt, Gerhard Jäger & Tonjes Veenstra (eds.), *Variation, Selection, Development*, 307–336. Berlin: Mouton de Gruyter.

Lai, Cecilia S. L., Simon E. Fisher, Jane A. Hurst, Faraneh Vargha-Khadem & Anthony P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523.

Lebeaux, David. 1989. Language acquisition and the form of the grammar. Amherst, MA: The University of Massachusetts dissertation.

Lightfoot, David. 1991. Subjacency and sex. *Language and Communication* 11, 67–69.

Locke, John L. & Barry Bogin. 2006. Language and life history: A new perspective on the evolution and development of linguistic communication. *Behavioral and Brain Sciences* 29, 259–325.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and Logical Form. *Linguistic Inquiry* 25, 609–665.

MacLean, Paul D. 1949. Psychosomatic disease and the 'visceral brain:' Recent developments bearing on the Papez theory of emotion. *Psychosomatic Medicine* 11, 338–353.

Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Nevins, Andrew. This volume. Two case studies in phonological universals: A view from artificial grammars.

Ouhalla, Jamal. 1991. *Functional Categories and Parametric Variation*. London: Routledge & Kegan Paul.

Perlmutter, David. 1978. Impersonal passive and the Unaccusative Hypothesis. *Berkeley Linguistics Society* 4, 159–189. University of California, Berkeley.

Piattelli-Palmarini, Massimo & Juan Uriagereka. 2005. The evolution of the narrow faculty of language: The skeptical view and a reasonable conjecture. *Lingue e Linguaggio* IV, 27–79.

Pinker, Steven & Michael T. Ullman. 2002. The past-tense debate: The past and future of the past tense. *Trends in Cognitive Sciences* 6, 456–463.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.

Platzack, Christer. 1990. A grammar without functional categories: A syntactic study of early child language. *Nordic Journal of Linguistics* 13, 107–126.

Poeppel, David. 2008. Linguistics and the future of the neurosciences. Paper presented at the DGfS workshop on *Foundations of Language Comparison: Human Universals as Constraints on Language Diversity*, Bamberg, Germany. (Universitat Bamberg, 27–29 February 2008.)

Poeppel, David & David Embick. 2005. Defining the relation between Linguistics and Neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.

Postal, Paul M. 1997. Islands. Unpublished manuscript, New York University.

Potts, Christopher & Tom Roeper. 2006. The narrowing acquisition path: From expressive small clauses to declaratives. In Progovac *et al.* (eds.), 183–201.

Progovac, Ljiljana. 2006. The syntax of non-sententials: Small clauses and phrases at the root. In Progovac *et al.* (eds.), 33–71.

Progovac, Ljiljana. 2008a. What use is half a clause? In Andrew Smith, Kenneth Smith & Ramon Ferrer i Cancho (eds.), *Evolution of Language: Proceedings of the 7th International Conference, Barcelona, Spain, March 12–15*, 259–266. Hackensack, NJ: World Scientific.

Progovac, Ljiljana. 2008b. Root small clauses with unaccusative verbs: A view from evolution. *Proceedings of FASL* 16, 359–373.

Progovac, Ljiljana. 2009a. Layering of grammar: Vestiges of proto-syntax in present-day languages. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 203–212. Oxford: Oxford University Press.

Progovac, Ljiljana. 2009b. Sex and syntax: Subjacency revisited. *Biolinguistics* 3, 305–336.

Progovac, Ljiljana. 2010a. When clauses refuse to be recursive: An evolutionary perspective. In Harry van der Hulst (ed.), *Recursion and Human Language*, 193–211. Berlin: Mouton de Gruyter.

Progovac, Ljiljana. 2010b. Imperative in compounds: Implications for historical and evolutionary studies. In Petr Karlík (ed.), *Development of Language through the Lens of Formal Linguistics*, 137–145. Munich: Lincom Europe.

Progovac, Ljiljana. To appear. Compounds and commands in the evolution of human language. *Theoria et Historia Scientiarum*.

Progovac, Ljiljana & John L. Locke. 2009. The urge to merge: Ritual insult and the evolution of syntax. *Biolinguistics* 3, 337–354.

Progovac, Ljiljana, Kate Paesani, Eugenia Casielles & Ellen Barton (eds.). 2006. *The Syntax of Non-Sententials: Multidisciplinary Perspectives*. Amsterdam: John Benjamins.

Pulvermüller, Friedemann. 2002. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press.

Pulvermüller, Friedemann. This volume. Brain–language research: Where is the progress?

Radford, Andrew. 1988. Small children's small clauses. *Transactions of the Philological Society* 86, 1–43.

Reis, Marga. 1995. Extractions from verb-second clauses in German? In Uli Lutz & Jürgen Pafel (eds.), *On Extraction and Extraposition in German*, 45–88. Amsterdam & Philadelphia, PA: John Benjamins.

Ridley, Mark. 1993. *Evolution*. Oxford: Blackwell Scientific Publications.

Roeper, Tom. 1999. Universal bilingualism. *Bilingualism: Language and Cognition* 2, 169–186.

Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT dissertation.

Schütze, T. Carson. 2001. On the nature of default case. *Syntax* 4, 205–238.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10, 80–126.

Stowell, Timothy. 1981. Origins of phrase structure. Cambridge, MA: MIT dissertation.

Stowell, Timothy. 1983. Subjects across categories. *The Linguistic Review* 2/3, 285–312.

Studdert-Kennedy, Michael. 1991. Language development from an evolutionary perspective. In Norman A. Krasnegor, Duane M. Rumbaugh, Richard L. Scheiefelbusch & Michael Studdert-Kennedy (eds.), *Biological and Behavioral Determinants of Language Development*, 5–28. Hillsdale, NJ: Erlbaum.

Stump, Gregory T. 1985. *The Semantic Variability of Absolute Constructions*. Dordrecht: Reidel.

Tchekhoff, Claude. 1979. From ergative to accusative in Tongan: An example of synchronic dynamics. In Frans Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*, 407–418. London: Academic Press.

Traugott, Elisabeth C. & Bernd Heine. 1991. *Approaches to Grammaticalization*, vol. II. Amsterdam: John Benjamins.

Ullman, Michael T. 2008. Variability and redundancy in the neurocognition of language. Paper presented at the DGfS workshop on *Foundations of Language Comparison: Human Universals as Constraints on Language Diversity*, Bamberg, Germany. (Universitat Bamberg, 27–29 February 2008.)

Uriagereka, Juan. 2008. *Syntactic Anchors: On Semantic Structuring*. Cambridge: Cambridge University Press.

Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Zeijlstra, Hedde. 2008. Hard and soft conditions on the Faculty of Language: Constituting parametric variation. *Linguistics in Potsdam* 28, 9–38.

*Ljiljana Progovac*
*Wayne State University*
*Department of English*
*Linguistics Program*
*5057 Woodward*
*Detroit, MI 48202*
*USA*
*progovac@wayne.edu*

# Brain-Language Research:
# Where is the Progress?

Friedemann Pulvermüller

Recent cognitive neuroscience research improved our understanding of *where*, *when*, *how,* and *why* language circuits emerge and activate in the human brain. *Where*: Regions crucial for very specific linguistic processes were delineated; phonetic features and fine semantic categories could be mapped onto specific sets of cortical areas. *When*: Brain correlates of phonological, syntactic and semantic processes were documented early on, suggesting language understanding in an instant (within 250 ms). *How*: New mechanistic network models mimicking structure and function of left-perisylvian language areas suggest that multimodal action-perception circuits — rather than separate modules for action and perception — carry the processing resources for language use and understanding. *Why* language circuits emerge in specific areas, become active at specific early time points and are connected in specific ways is best addressed in light of neuroscience principles governing neuronal activation, correlation learning, and, critically, partly predetermined structural information wired into connections between cortical neurons and areas.

## 1.    Introduction: Questions in Focus

The aim of the neuroscience of language is to find the brain correlates of linguistic processes and representations. Correlates of linguistic representations are sought in neuronal structures, that is, nerve cell circuits, and correlates of linguistic processes are sought in patterns of neuronal activation. These aims have as yet not been reached. In many cases, conclusions are still at the level of 'areas' 'performing' certain functions, a state not untypical for cognitive neuroscience in general. However, such 'arealogy' can be understood as an intermediate step on the journey towards neuroscientific explanation. To keep the ultimate destination in sight and in focus, it may be relevant to pause and check.

Ultimately, the clarification of the brain correlates of a given cognitive representation R and process P implies answers to (at least) four critical questions:

1.  *Where*-question: Which brain parts, areas, and, eventually, neurons are active during, and are critical for, process P and the representation(s) R P relies on?

2.  *When*-question: At which point in time in the usage or understanding of language does process P occur; when is representation R activated and processed?

3.  *How*-question: Which neuronal circuit, which nerve cells linked in which way, is the brain basis for representation R; which spatiotemporal pattern of neuronal activation in this circuit does underpin the process P?

4.  *Why*-question: For what reason are R and P located in these specific brain parts and activated at these specific points in time, and why is R laid down in this specific neuronal circuit, P being expressed by these specific activation patterns?

The present contribution will briefly review research addressing critical facets of these four questions. A focus will be on recent progress in mapping *specific* linguistic representations and processes onto brain space and time and a second focus will be on circuit structure and function.

## 2.    *Where*? Mapping Modules — Mapping Features and (Sub-)Categories

### 2.1.    Where-*Question: Meaning*

Once, the name of the game in the cognitive neuroscience of language was to find a place in the brain for the major modules of linguistic processing. For example, when I wrote a paper for the journal *Behavioural and Brain Sciences* in 1999, a number of colleagues, well-known leaders in the field, commented on my review of the cortical basis of semantic processing, most of them by communicating that they had good empirical evidence to believe that a specific brain part is particularly relevant for word meaning processing (see comments, Pulvermüller 1999). This 'meaning centre' as one may want to dub it, was placed in different parts of the left hemisphere, so that a large part of the left hemisphere was covered with semantic areas and reconciling the different views with each other appeared difficult. A more recent update of the literature shows a similar picture, especially in the temporal lobe, on which many studies focus. For example, Hickok & Poeppel (2007) put their 'lexical interface' assumed to connect phonological and semantic representations in the middle-temporal cortex, Scott & Johnsrude (2003) suggest the anterior part of the superior-temporal gyrus as the meaning interface, and Patterson *et al*. (2007) put — based on a wealth of evidence from degenerative brain disease —that the temporal pole is the key area for semantic processing. Figure 1 illustrates the variability of positions.
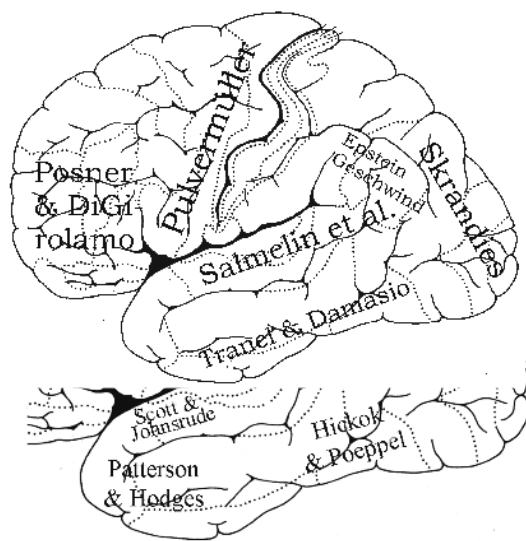
*Figure 1: Different authors view different cortical sites as most important for binding the meaning of words to their form. The sketched proposals are based on recent publications (Epstein 1999, Posner & DiGirolamo 1999, Pulvermüller 1999, Salmelin et al. 1999, Skrandies 1999, Tranel & Damasio 1999, Scott & Johnsrude 2003, Hickok & Poeppel 2007, Hodges & Patterson 2007). This paper attempts at developing an integrated perspective.*

The principal problem of the debate about such *unitary* meaning centers — or centers whose function it is to bind any meaning to any word/symbols — is the following: There is solid evidence for the importance of various cortical areas, at least in temporal and frontal lobes, in semantic processing and, by accumulating more evidence in favor of the importance of any one area, one cannot, evidently, disprove the role of the other ones. This would only be the case if there was an exclusive either–or, that is, if only one cortical area was allowed to include a major meaning switchboard. Although semantic processing implies the integration of information from different sensory modalities, such integration can be computed locally between adjacent neurons as it can be carried by distributed populations of interacting neurons; hence again no need for a unitary semantic area.

A second major problem for a unitary meaning approach is semantic category-specificity: Lesions in many cases do not affect all word kinds (and symbol types) to the same degree. Dependent on where the lesion is situated, specific categories of knowledge are affected more or less. Significant differences between semantic kinds — such as animal vs. tool names — have been documented with lesions in frontal and temporal cortex (Warrington & McCarthy 1983, Gainotti 2006) and processing differences between fine-grained semantic categories have even been reported in patients with semantic dementia (Pulvermüller *et al.*, in press). Here, the solution lies in the integration of general lexico-semantic and category-specific semantic processes, as may be manifest in the interaction between a range of cortical areas (Patterson *et al.* 2007, Pulvermüller *et al.*, in press). The precise location — or, perhaps better: distribution — of general and category-specific semantic circuits is one of the hottest topics in current neuroscience research. Figure 2 shows recent data indicating the approximate lo-

cations of category-specific semantic circuits, as they can be inferred today from neuroimaging data, and contrasts them with brain activations generally seen for meaningful written word stimuli.
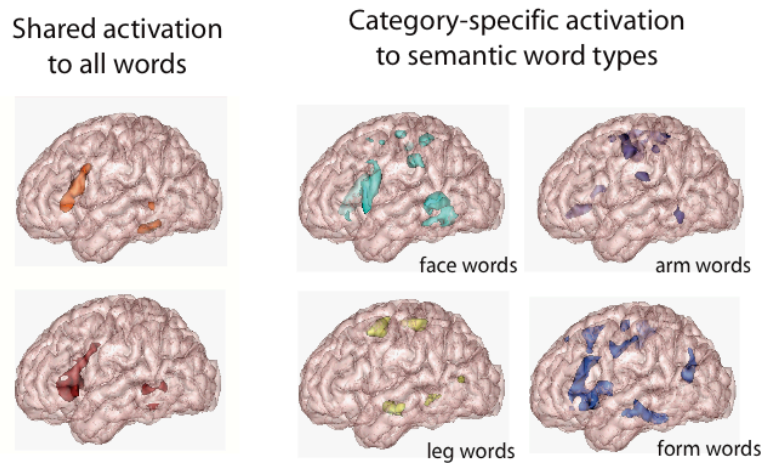


*Figure 2: Brain activation patterns during passive word reading: Cortical areas activated by all words alike (left side) are contrasted with areas specifically activated by fine-grained semantic word categories (right side), action words related to the face (lick), arms (pick) or legs (kick) and visually-related form words (square) (modified from Pulvermüller, Kherif et al. 2009). Areas found active generally to all kinds of words may indicate the distribution of circuits for processing of general lexical-semantic information, whereas the widely distributed area sets found active for specific semantic types may index the distribution of category-specific semantic circuits.*

So where is the progress? It still lies in the mapping of meaning on brain matter. Not just in the mapping of *any kind of meaning* to brain structure, or the delineation of a unitary meaning centre, global semantic binding site or the like, but in the brain mapping of sometimes fine grained semantic categories and sub-types of knowledge. Most words indeed activate middle and inferior-temporal areas mainly involved in the processing of visual information about objects. This is not surprising because most words in languages like English are nouns refer-ring to objects known through the visual modality. Animal and tool words, and similarly their related concepts, activate different inferior-temporal and middle-temporal areas in both hemispheres (Damasio *et al.* 1996, Chao *et al.* 1999, Martin 2007), and words referring to objects with characteristic form or color features (square vs. coal) elicit activity in overlapping but distinct areas in bursiform, pa-rahippocampal and middle temporal gyri (Moscoso Del Prado *et al.* 2006, Pulver-müller & Hauk 2006, Simmons *et al.* 2007). The inferior-temporal cortex — from pole to temporo-occipital junction — reflects a range of semantic distinctions and lesion in this region also appears to lead to specific degradation of particular semantic categories, to category specific semantic deficits (Warrington & Shallice 1984, Damasio *et al.* 1996, Miceli *et al.* 2001, Neininger & Pulvermüller 2003). As one example, lesion of rostro-mesial temporal cortex in the left hemisphere — a subpart of which (anterior parahippocampal gyrus) was found active specifically during color word processing — impairs object color knowledge specifically (Miceli *et al.* 2001).

The temporal lobes are not the only key areas for semantic processing. Words loaded with affective-emotional meaning can activate the amygdale, insular structures, and the posterior cingulated cortex (Straube *et al.* 2004, de Araujo *et al.* 2005). Odor words, as compared with matched control words, activate olfactory cortex along with limbic structures (Gonzalez *et al.* 2006), sound related words activate the superior temporal lobes more strongly than matched control words (Kiefer *et al.* 2008) and, critically, action-related verbs spark the motor and premotor cortex in such a specific manner that the body part related-ness of the action indexed by the words becomes manifest in somatotopic acti-vation in the motor strip (Pulvermüller *et al.* 2000, Hauk *et al.* 2004, Shtyrov *et al.* 2004, Pulvermüller, Shtyrov & Ilmoniemi 2005). Words such as 'pick' and 'kick' would therefore specifically activate areas also active when subjects move their finger or foot. The overlap between areas active during motor performance and during congruous word processing is not complete; notably, normal motor per-formance creates somatosensory input, leading to somatosensory postcentral activation which, when overlaid with motor cortex activation, shifts the centre of gravity of activation backward, towards the parietal lobe. However, the somato-topic line-up of premotor activity reflecting aspects of action semantics could be replicated by a range of studies (Tettamanti *et al.* 2005, Aziz-Zadeh *et al.* 2006, Tomasino *et al.* 2007, Kemmerer & Gonzalez-Castillo 2010, Boulenger *et al.* 2009, Raposo *et al.* 2009), with occasional failure to replicate activity in specific regions of interest (Postle *et al.* 2008). In one study, the semantic somatotopy could even be documented in abstract idiom processing ('grasp the idea', 'kick the habit'; cf. Boulenger *et al.* 2009) consistent with an embodied, partly compositional view on abstract sentence meaning construction, to which lexical meaning contributes (Lakoff 1987, Barsalou 1999).

Importantly, these motor activations seem to index critical parts of the cortical semantic processor. Lesions in the motor system impair the processing of action-related words, especially that of action verbs (Damasio & Tranel 1993, Daniele *et al.* 1994, Neininger & Pulvermüller 2003, Tranel *et al.* 2003, Gainotti 2008) and, in addition to these, of the related action concepts (Bak *et al.* 2001, Bak *et al.* 2006). In healthy individuals, magnetic stimulation below the motor thresh-old to hand and foot areas in the left motor cortex could be shown to facilitate the processing of hand and foot related words specifically (Pulvermüller, Hauk *et al.* 2005). These results document a causal role of the motor system in processing action concepts and words semantically related to actions.

What we have learned is, therefore, that the level of specificity of brain-meaning mapping is much greater than previously thought. This is exciting from a linguistic perspective, as some semantic features of words seem to be apparent from the brain response they elicit. Of theoretical importance here is the fact that semantic areas could be predicted *a priori* on the basis of brain-theory, lending strong evidence for the underlying explanatory model (see section 5 below in the *why*-section of this article). Very specific action and perception features of refer-ential semantic information linked to words can be mapped onto cortex. The search for the unitary meaning centre has, however, led to much disagreement, although it is possible that meaning integration at highly abstract levels draws upon only one area. The meaning centre seems to be best described as the union

of brain areas critically involved in category-specific processing and the bias to-
wards temporal cortex may relate to the habit of researchers to test object nouns
and their related concepts. Note that important knowledge about most objects
comes through the visual modality and the involvement of the inferiortemporal
stream of object processing is therefore not surprising. Even for abstract words
and sentences, different areas were found active by different researchers (e.g.,
Noppeney & Price 2004, Binder *et al.* 2005, Boulenger *et al.* 2009) raising the
question whether category-specificity might hold even at abstract semantic levels
(Pulvermüller & Hauk 2006). In one view, gradually more abstract semantic re-
presentations develop in progressively anterior areas in temporal and frontal
cortex as a consequence of sensorimotor activity (Pulvermüller 2008).

An integrated view proposes category-specific semantic circuits whose
precise distribution depends on meaning type (cf. Fig. 2). Areas most important
for meaning emerge close to left-perisylvian language cortex — especially the
inferior-frontal and superior-temporal gyri and sulci along with the underlying
insula. All linguistic functions depend on this perisylvian region, whereas the
category-specific meaning circuits extend throughout the cortex, the extrasylvian
space. Action and object related meaning circuits draw upon motor and sensory
areas and abstract semantic circuits develop in the vicinity of these sensorimotor
sites, in anterior temporal and prefrontal cortex. There is differential laterality of
linguistic and semantic processes and representations. Due to some property of
the left perisylvian cortex (see the *why*-section 5), linguistic circuits are generally
lateralized, although semantic circuits are spread out more symmetrically
throughout both hemispheres (Fig. 2; Pulvermüller & Mohr 1996, Pulvermüller,
Kherif *et al.* 2009).

Although the recent support for category-specific semantic circuits appears
as a milestone in understanding the brain basis of meaning, it should not be
ignored that some colleagues expressed criticisms. Caramazza's group suggested
that motor activity during the processing of action verbs may not be related to
semantic processes but may instead be an epiphenomenon related to mental
images being retrieved, if not entirely irrelevant 'overflow' activation (Oliveri *et
al.* 2004). In face of more recent neuropsychological evidence supporting a crucial
role of motor systems for processing words of specific action-related semantic
categories (for review, see Pulvermüller & Fadiga 2010), a new proposal now ack-
nowledges a (possible) semantic function of the motor system, but complements
it with an abstract symbol processor (Mahon & Caramazza 2008), a view similar
to Patterson *et al.*'s (2007) suggestion that a 'semantic hub' — according to their
data, in the temporal pole — complements widely distributed category-specific
semantic circuits (see also Pulvermüller *et al.*, in press).

A common misunderstanding about the role of sensorimotor circuits in
semantic processing is that they provide the *only* source of meaning knowledge.
However, this position does not appear very plausible. Combinatorial knowledge
about words regularly occurring in sentence and discourse contexts implies
semantic knowledge, for example about the most frequent color word the item
'strawberry' would co-occur with (Landauer & Dumais 1997). Combinatorial
word properties allow not only the classification of words into syntactic classes,
they also lead to distinctions along semantic boundaries, separating types of ob-

jects and types of actions (Pulvermüller & Knoblauch 2009). A mechanistic neuro-biological approach captures the storage of the underlying word–word correlations by way of the very same mechanisms it also uses for storing word–world correlations in neuronal links between sensorimotor and perisylvian language cortices (Pulvermüller 2010). Furthermore, correlation learning is not restricted to the single word level, but can, in principle, occur for larger constructions, especially if they are being used stereotypically in specific contexts (Goldberg 2003). Current neuroimaging results seem consistent with a contribution of semantic representations of both constituent words and whole constructions when the meaning of abstract idiomatic sentences is being processed (Boulenger *et al.* 2009).

Some issues in the cortical localization of semantic processes are still open. The idea that access to movement knowledge tied to words is reflected in lateral temporal activation just anterior to a movement sensitive visual processing area (Martin *et al.* 1995) was recently questioned based on a lack of activation differences between nouns with more or less semantic relationship to movement (Bedny *et al.* 2008). While this finding argues against a role of middle temporal cortex in kinematic semantics, there is still solid evidence that the action-relatedness of word meaning is reflected in the activation of the left middle temporal area (MNI coordinates –62/–52/4; Hauk *et al.* 2008). The fact that the area activates more strongly to verbs than for nouns (e.g., Bedny *et al.* 2008, Hauk *et al.* 2008) is consistent with the action relatedness of most verbs, even verbs used to speak about so-called 'internal states'. States such as thinking and feeling have characteristic behavioural expressions, thus intrinsically linking the semantics of the respective terms to action (Wittgenstein 1953). Therefore, any noun-verb difference is hopelessly confounded with semantic differences (Pulvermüller *et al.* 1999). Furthermore, a recent study suggested that in the left middle temporal area, there are, side by side, different subareas that respond to words generally (–53/–49/–1), thus possibly contributing to general lexico-semantic processes, and to very specific semantic subcategories of action verbs (e.g., hand-related action verbs, –49/–51/–9) (Pulvermüller, Kherif *et al.* 2009). Such fine subcategorization may be a consequence of recurrent connections with the motor system, where semantic somatotopic activation is established. If the middle-temporal activation to action-related words is due to links with the motor system (rather than to knowledge about moving visual input), it becomes explainable why such activation persists in visually deprived individuals (Mahon *et al.* 2009) who are not principally limited in their action repertoire and typically learn words, even visually-related ones, in action contexts (see Landau & Gleitman 1985). These data are consistent with a differential role of temporal and frontal areas in semantic processing, although more research may indeed help clarifying the various linguistic roles of middle temporal gyrus activation in word and sentence processing.

## 2.2.  Where-*Question: Speech Sounds*

Phonological processes are located in perisylvian cortex. In one view, speech analysis is attributed to systems in the anterior-lateral (antero-ventral processing

steam) and/or posterior part of the superior-temporal cortex (postero-dorsal stream, including planum temporale and lateral superior-temporal gyrus) (Rauschecker & Scott 2009). The planum temporale and other posterior superior-temporal areas have long been viewed as critical for language perception and understanding, based on evidence from clinical language deficits (see, e.g., Geschwind 1970). Recent neuroimaging experiments showed that speech yields stronger activation in antero-ventral superior-temporal areas compared with matched noise patterns (Scott *et al.* 2000, Uppenkamp *et al.* 2006), and this evidence is also consistent with data from macaques that anterior superior-temporal activity indexes species-specific calls (Romanski *et al.* 1999). Similar responses in posterior superior-temporal cortex to speech and other acoustic stimuli still allow for a role of this region in speech-language processing. This observation is compatible with a view of postero-dorsal areas, especially planum temporale but possibly also temporo-parietal junction, as a 'computational hub' for processing spectrotemporally rich acoustic patterns (Griffiths & Warren 2002). In addition to superior-temporal cortex, inferior-frontal cortex is active during listening to speech, as could be demonstrated using TMS (Fadiga *et al.* 2002), and inferior-frontal activation even persists during passive exposure to speech, as could be shown using MEG (Pulvermüller 2003, Pulvermüller, Shtyrov & Ilmoniemi 2003). Critical inferior-frontal areas include posterior Broca's (pars opercularis) and premotor cortex (Wilson *et al.* 2004). Similar to the posterior superior-temporal cortex, the motor system's role is not confined to speech processing. The sounds of actions activate different sections of the fronto-central sensorimotor cortex in a very similar manner as linguistic sounds do (Hauk, Shtyrov & Pulvermüller 2006, Lahav *et al.* 2007). These results suggest that the computational hub for sound processing extends from posterior-temporal cortex to inferior-frontal and premotor regions. Precisely timed spatio-temporal patterns of cortical activation spreading in this distributed cortical system may signify the processing of speech and other action sounds (Pulvermüller & Shtyrov 2009).

Similar to the semantic domain, recent advances in our knowledge about phonological representations and processes in the brain relates to specificity. Different areas in superior-temporal cortex were found active when subjects listened to different kinds of speech sounds (Diesch *et al.* 1996, Obleser *et al.* 2003, Obleser *et al.*, 2006, Pulvermüller *et al.* 2006, Obleser *et al.* 2007). Typical examples of the phonemes [p] and [t] for example were mapped to adjacent areas in superior-temporal gyrus, anterior to primary auditory cortex and Heschl's gyrus. Interestingly, a similar phonological mapping was evident in the motor system, where the production of [p] and [t] activated different precentral areas in a soma-totopic fashion. The articulatory mapping of phonemes to the motor system corresponded to the localization of the articulators mainly involved in the production of the respective speech sounds — the lips for [p] and the tongue for [t] (Lotze *et al.* 2000, Hesselmann *et al.* 2004).

Notably, these different precentral motor/premotor areas were also found active during listening to speech. Listening to [t] activated the precentral focus also excited when producing a [t] or moving the tongue tip, and when hearing [p], a slightly dorsal area also active when producing this phoneme or when moving the lips lighted up (Pulvermüller *et al.* 2006). The critical role of these

motor systems in speech perception is evident from TMS work stimulating the motor regions of the lips and the tongue: Such stimulation biases the speech comprehension system in favor of congruent sounds. Therefore, when the tongue (or lips) area was stimulated, subjects tended to perceive [t] (or [p]) sounds more quickly, or even to misperceive [p] sounds as [t] (or the reverse) (D'Ausilio *et al.* 2009). This observation demonstrates that motor systems critically contribute to the speech perception process.

In the phonological domain, progress seems two-fold. First, the perisylvian cortex, which is well-known to be critical for phonological processing and representation, can be further subdivided according to phonological properties. Phonetic distinctive features, DFs, and speech sounds discriminated by these DFs can be mapped on different brain substrates in inferior-frontal and (antero-lateral) superior-temporal cortex. Second, the temporal and frontal neuronal ensembles appear to interact with each other and to be functionally interdependent in phonological processing. The summarized data argue against proposals that seem to play down the role of frontal cortex in speech perception (see the *how*-section 4 below, Hickok & Poeppel 2007, Lotto *et al.* 2009, and Scott *et al.* 2009). Note again that left inferior-frontal cortex activates in speech perception even when subjects try to ignore incoming speech sounds. Frontal activation therefore does not depend on attention being focused on speech (Pulvermüller *et al.* 2003, Pulvermüller & Shtyrov 2006, 2009), although attention certainly exerts a modulatory function on language-elicited brain activity (Garagnani, Shtyrov & Pulvermüller 2009, Shtyrov *et al.*, in press). In the language domain, the posterior-dorsal vs. anterior-ventral stream debate seems, at present, not fully conclusive, as both parts of the superior-temporal cortex are apparently involved in speech processing and the absence of phoneme specificity in the posterior superior-temporal cortex appears as a null result without strong implications. Clear evidence exists for anterior-lateral superior-temporal activation discriminating phonemes from noise and phonemes between each other, but a contribution of posterior parts of superior-temporal cortex to phonological processing also receives support.

### 2.3.    Where-*Question: Syntax*

It would seem exciting to delineate cortical maps for rules of syntax, similar to the mapping of semantic categories and that of phonetic DFs reviewed earlier in this section. However, such syntactic mapping has so far not been fully successful and major reasons for this lie in the tremendous difficulties the grammar domain creates for the experimental scientist. When comparing grammatical sentences to word strings with syntactic errors, the latter elicit stronger brain activation in left perisylvian cortex, especially in inferior-frontal and in superior-temporal cortex (e.g., Friederici *et al.* 2000, Indefrey *et al.* 2001, Pulvermüller & Shtyrov 2006, and Friederici 2009). When directly comparing sentences with different grammatical structure, for example active and passive, subject and object relative, and coordinated and subordinated sentences, the grammatically more demanding sentences tended to elicit stronger activation; again some of the activation differences were located in left perisylvian cortex (Just *et al.* 1996, Caplan *et al.* 2000, Caplan *et al.* 2008). Although these results suggest that processing of grammaticality and of

the complexity of grammatical structure relates to inferior-frontal and superior-temporal circuits, they do not unambiguously prove this. Ungrammatical sentences are rare and therefore exceptional, whereas grammatical ones are normally more common, and among the grammatical ones, the sentences considered to be more complex (e.g., object relatives) are rarer than the ones considered to be simpler (e.g., subject relatives). Heroic attempts have been made to control sequential probabilities while, at the same time, varying grammatical structure of well-formed sentences (e.g., Bornkessel *et al.* 2002). However, as has been argued by linguists, such control has not been perfect (Kempen & Harbusch 2003) making it seem impossible to exclude the probability confound.

Whether syntax depends on discrete combinatorial rules or is best described in terms of sequential probabilities constitutes a major debate in cognitive science (McClelland & Patterson 2002, Pinker & Ullman 2002). The linguistic position that rules of syntax and universal underlying principles govern grammar is in contrast with approaches using systematic probability mapping in neural networks or statistical procedures lacking any rule-like symbolic representations. These, too, are capable of modeling linguistic processes and have the additional advantage of explaining aspects of the learning of grammar (Rumelhart & McClelland 1987, Hare *et al.* 1995). That certain types of syntactic (and, likewise, phonological, semantic) structures *a priori* require a system of discrete combinatorial rules and representations may, therefore, appear as a too strong statement, although this assumption figures as a firm corner stone of much cognitive theorizing in the second half of the 20th century. Whether discrete representations and especially rule-like entities exist turns out to be an empirical issue but also one addressable by brain theory and brain-based modeling (see *how*-section 4). The term 'discrete' — an expression with many facets that is used in different areas of cognitive science with rather different meanings — is used here to refer to a mechanism that either is being engaged in a given condition or is not, with little if any room for gradual intermediate steps. The sentence 'Build a sentence from (at least) a noun and a verb' describes a discrete combinatorial mechanism at an abstract linguistic level. Can we expect that such discrete rule-like processes, rather than probability mapping, are effective at the neurobiological level?

Empirical testing of the existence of rule-like mechanisms is possible if probability mapping and rule-applicability dissociate. Examples are sentences that are grammatically correct but extremely rare in language use. These can be contrasted with grammatical sentences that are common, but also with ungrammatical strings that are rare to a similar degree as the rare grammatical items. As mentioned, ungrammatical strings elicit stronger brain activity than common grammatical sentences. As this neurophysiological difference is even observable if the same, identical recordings of spoken word strings are presented many times and even when subjects do not pay attention to the speech stimuli, some grammatical brain processes appear to be automatic (Shtyrov *et al.* 2003, Hasting & Kotz 2008, Pulvermüller *et al.* 2008). But, critically, would a rare gram-matical string produce a brain response indistinguishable from that of a common grammatical string, as a discrete all-or-nothing approach might suggest, or would the gradual probability differences between the strings be reflected in the neurophysiological brain response, as a probability mapping theory would predict? The

former was found: The brain response to rare ungrammatical strings was enhanced, that to grammatical strings was attenuated, regardless of the sequential probability of their constituent words (Pulvermüller & Assadollahi 2007). This result pattern is consistent with, and therefore supports, the rule theory. It should however be noted that a neural approach without discrete representations can be modified to fit these data if appropriate non-linearities are built into the network.

So, again, where is the progress? In the delineation of category-specific semantic circuits distributed over specific sets of cortical areas, in the mapping of phonetic features onto brain systems that encompass superior-temporal and inferior-frontal (including premotor) areas, and in new evidence in favor of discrete combinatorial rules brain-supported by left perisylvian circuits.

## 3.      *When*? The Rapid Time Course of Language Understanding

A main stream view held that language understanding is a relatively late process (for review, see Barber & Kutas 2007 and Pulvermüller *et al.*, 2009). Semantic processing, along with lexical ones, were assumed to be first indexed by the N400 component of the event-related brain potential and field. Syntactic processing was assumed to be indexed by an even later component, called P600. Both components peak around half a second after information necessary for identifying critical stimulus words is present, suggesting that at least this amount of time elapses between presence of a word in the input (say, 'bear' in a warning context), and the initiation of an appropriate response (for example, running away). Such long-delay comprehension systems may have advantages under certain conditions, however, from a Darwinian perspective, a faster system minimizing the comprehension latency would certainly have constituted an evolutionary advantage.

In fact, some research indicated early brain reflections of syntactic and semantic processing. In the semantic domain, the meaning of action words becomes manifest in somatotopic motor systems activation already 100–250 ms after availability of information about the identity of spoken (Shtyrov *et al.* 2004, Pulvermüller, Shtyrov & Ilmoniemi 2005) or written (Pulvermüller *et al.* 2000, Hauk & Pulvermüller 2004) stimulus words. At the same latencies, the neuro-physiological responses dissociated between word kinds semantically linked to visual information — for example, between color- and form-related words (Sim & Kiefer 2005, Moscoso Del Prado Martin *et al.* 2006, Kiefer *et al.* 2007). Likewise, large word categories differing in both their grammatical function and their semantic characteristics — for example, grammatical function words and referential content words, or object nouns and action verbs — dissociate neurophysiologically within 250 ms (for an overview, see Pulvermüller, Shtyrov & Hauk 2009). A range of other psychological and linguistic factors, including frequency of occurrence of words and their parts and general semantic properties, could also be mapped onto the first 250 ms, in this case calculated from the *onset* of written word presentation. The early effects (<250 ms) seem to be more variable and less robust than the later (~500 ms) ones (Barber & Kutas 2007). Importantly, they do depend strongly on stimulus properties, especially the length and lumi-

nance of written words and the loudness of spoken works and the point in time when they can first be recognized with confidence (Pulvermüller & Shtyrov 2006). Early and late indexes of semantics may reflect different processes in the analysis of word meaning, automatic semantic access and semantic re-analysis (Pulvermüller, Shtyrov & Hauk 2009).

Syntactic processing is long known to have an early brain correlate. Violations of phrase structure rules were found to lead to enhanced negativities already at 100–250 ms, in the early left-anterior negativity (Neville *et al.* 1991, Friederici *et al.* 1993). More recently, similar early violation responses, in the syntactic mismatch negativity component, have been reported to violations of the rules of phrase structure and agreement (Shtyrov *et al.* 2003, Hasting & Kotz 2008, Pulvermüller *et al.* 2008). It is these early responses that are automatic (see *where*-section above), whereas the late ones (P600) depend on attention to stimulus sentences: Brain correlate of syntactic mismatches in the 'syntactic mismatch negativity' can be recorded in subjects who do not attend to speech that includes grammatical violations; the early responses (up to 150 ms) remain unchanged even if subjects are heavily distracted from speech and syntax by a continuously applied attentional streaming task. This proves that at least some early mechanism of the brain's grammar machinery operates automatically or, as once claimed, 'like a reflex' (Fodor 1983).

In sum, syntactic and semantic processes are reflected by relatively late brain responses that depend on task and attention to stimuli. In addition, early, <250 ms, brain indexes of syntax and semantics also exist and these seem to be less dependent on attention being paid to stimuli, in some cases entirely attention-independent. For phonological and pragmatic processes, there are also reports about early as well as late brain correlates. For example, an early brain correlate of phoneme processing is present in the mismatch negativity (Dehaene-Lambertz 1997, Näätänen *et al.* 1997) and early indexes of pragmatic deviance have been reported in studies of text processing (Brown *et al.* 2000). Phonological expectancy violations (Rugg 1984, Praamstra & Stegeman 1993) as well as pragmatic and discourse-related ones (van Berkum, Brown *et al.* 2003, van Berkum, Zwitserlood *et al.* 2003) also were found to produce late effects at ~400 ms. An interpretation of early vs. late effects is possible along the lines of dual-stage models, such as Friederici's (2002) influential model. The early process would accordingly be an automatic comprehension or matching process, whereas the late process could either imply an in depth extension of the process or a revision and re-analysis. Friederici proposes this for the syntactic domain but, in light of the early and late components indexing essentially all kinds of psycholinguistic information, the same general concept can be applied to other psycholinguistic levels of processing, too.

A different and complementary approach relates the latency of cognitive neurophysiological responses to stimulus properties, especially the variability of physical, form-related properties. Larger variance of these variables, including word length and acoustic properties of spoken materials, increases the variance of brain responses especially at early latencies. Therefore, such variance may mask early brain responses reflecting cognitive and psycholinguistic processing. Late responses survive in attention-demanding tasks because they are large,

long-lasting and widespread. Evidence for this view has recently been reported. Minimal variability of physical and psycholinguistic stimulus properties is critical for obtaining early effects in lexical, semantic, and syntactic processing (Assadollahi & Pulvermüller 2001, Pulvermüller & Shtyrov 2006, Penolazzi *et al.* 2007).

Important time aspects are immanent to the orchestration of cortical sources. It has been suggested that the brain uses a precise temporal code to transmit information at the neural level. Serial models had put that phonological, lexical, syntactic, and semantic processes follow each other in a given order (which varied between models). However, the early near-simultaneous neurophysiological responses mentioned above suggest that these processes run largely in parallel with little if any offset between them (Hauk, Davis *et al.* 2006). Furthermore, small timing differences have been reported between subtypes of semantic processes. Leg-related action words tend to spark the leg region slightly later than words with arm or face reference activate their corresponding inferior motor areas (Pulvermüller, Shtyrov & Ilmoniemi 2005). Here, timing differences seem to indicate semantic differences. A recent study comparing the timing of left superior-temporal, lateral-central, and inferior-frontal area activations found delays between regions that depended on stimulus type. Phoneme sequences and word stimuli led to a delayed activation of inferior-frontal cortex, 10–25 ms after superior-temporal cortex, whereas noise stimuli failed to elicit a comparable activation delay between regions (Pulvermüller & Shtyrov 2009). Therefore, the delay between regions of interest activations coded the phonological status of acoustic stimuli.

Interestingly, reliable time delays only emerged between superior-temporal and inferior-frontal cortex. The latero-central region including motor and premotor areas activated together with superior-temporal cortex. This suggests that the postero-dorsal stream activates more quickly than the antero-ventral stream, but that the latter conveys important information about the phonological status of sounds.

In summary, early near-simultaneous brain responses (latency <250 ms) index different facets of the comprehension process, including word form analysis, semantic access along with syntactic and semantic context integration, suggesting near-simultaneity (or short-delay seriality) in psycholinguistic information access. The short delays are potentially accountable in terms of cortical conduction times (Pulvermüller *et al.* 2009).

## 4.   *How*? Brain-Based Models of Circuits, Their Activations and Delays

A general conclusion suggested by much recent research is that action and perception are not stand alone processes but are functionally interwoven at the mechanistic level of neuronal circuits. This insight has gained momentum in basic and cognitive neuroscience, including research into perception and action in animals and humans (Rizzolatti & Craighero 2004, Rizzolatti & Sinigaglia 2010), visuomotor integration (Bruce & Goldberg 1985), and language processing (Pulvermüller 2005, Pulvermüller & Fadiga 2010). Consistent with such interdependence are both behavioural and neurofunctional observations. Action does not

solely relate to motor systems activation, but likewise draws on perceptual mechanisms, as in the co-activation of superior-temporal cortex during silent speaking (Paus *et al.* 1996). The Lee effect of delayed auditory feedback on speech output (Lee 1950) demonstrates a profound automatic influence of acoustic-phonetic processes on ongoing speech output, a conclusion strengthened by the neuro-functional studies demonstrating motor systems activation by perception of movements and speech (Fadiga *et al.* 1995, Fadiga *et al.* 2002). Speech perception does not just 'take place' in the superior-temporal cortex alone. The motor system is co-activated and assists, modulates, and sharpens the speech perception process.

One may describe this interaction in classic terms. Earlier proposals had suggested that perception is, in part, an active process by which action hypotheses play a role (see also Bever & Poeppel, this volume). Accordingly, 'bottom-up' perceptual analysis triggers a hypothesis about the input, followed by an action-related 'top-down' synthesis, the product of which is finally compared with and eventually matched to further input information, thus confirming or rejecting the perceptual hypothesis (Halle & Stevens 1959, 1962). When recognizing a naturally spoken syllable such as [pIk], already the vowel includes co-articulatory information about the subsequent consonant, which may give rise to the — still premature — hypothesis that [pIk] is coming up (Warren & Marslen-Wilson 1987, 1988). This hypothesis can be compared with further input — especially the plosion of the final [k] — until a match is reached. In the syntactic domain, a noun may generate the hypothesis that a sentence including a verb will emerge, and the N–V hypothesis can be compared, and eventually matched, with the input, a process compatible with the state sequence of a left-corner parser (Aho *et al.* 1987). This hypothesis generation and testing, as suggested by classic cognitive theories of phonological analysis-by-synthesis and left-corner parsing may capture aspects of the functional significance of the co-activation of frontal and posterior areas in cognitive processing in general and in language processing in particular.

While it is possible that similar descriptions in terms of perception-by-synthesis may capture aspects of the real mechanisms, they are not very precise and certainly not spelt out in terms of neurons. A vague formulation of action-perception interaction still leaves open questions including the following: How many concurrent hypotheses can be entertained at a time, and how many simultaneous top down predictions are allowed? Can analysis, hypothesizing, and synthesis, and matching run in parallel, with constant functional interaction between them, or are they serial modular processes? Are controlled conscious intentional decisions required in the literal sense or can the 'decision' process also be construed as automatic? There are many degrees of freedom here. One kind of description — in terms of 'hypotheses' and 'synthesis' — suggests attention demanding modular processes that are being entertained sequentially, one by one. However, much psycholinguistic research supports parallel processing of competing hypotheses. Gating experiments for example indicate that several competing hypotheses about possibly upcoming words are built, maintained and tested in parallel until one of them 'wins', a position immanent to models in the tradition of the cohort theory (Marslen-Wilson 1987, Gaskell & Marslen-Wilson 2002). The motor theory of speech perception, as one variant of an analysis-by-

synthesis approach, made the additional assumption that speech perception as a modular process is separate from other acoustic perceptual processes, a view difficult to reconcile with current neuroimaging data showing that the same cortical foci are active when producing articulator movements and speech sounds (Pulvermüller *et al.* 2006). An obvious deficit of analysis-by-synthesis approaches is the lack of a time scale. Perceptual analysis, hypothesis generation, and action synthesis and matching could each last for seconds or take place near-simultaneously. This general approach seems to be in need of additional detail to provide mechanistic explanations.

In my view, progress in clarifying the mechanisms underlying action-perception interaction requires brain theory. Correlational information between the speech signal and articulatory gestures together with neuroanatomical and neurophysiological knowledge and principles provide a firm basis for postulating action-perception circuits that (i) span inferior-frontal and superior-temporal areas, (ii) become active near-simultaneously with minimal inter-area delays determined by axonal conduction times, (iii) play a role in speech production and in speech perception too, and (iv) provide continuous facilitatory interaction between the inferior-frontal and superior-temporal parts of each circuit while at the same time (v) competing with other action-perception circuits.

The mechanisms of and processes in such circuits can be explored in computational work using networks with neuroanatomically realistic structure and plausible neurophysiological function (see Fig. 3; Wennekers *et al.* 2006, Garagnani *et al.* 2008).
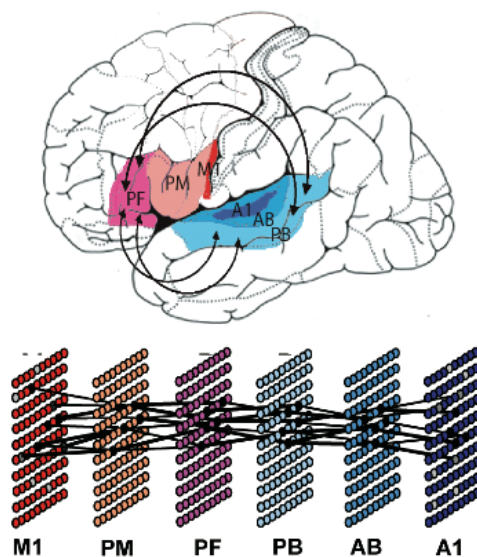


*Figure 3:    Areas of the left-perisylvian language cortex, connections between them and implementation in the model of the language cortex (MLC; Garagnani et al. 2007, Garagnani et al. 2008). Explicit neuromechanistic models grounded in neuroanatomy and neurophysiology can be used to simulate language processes in the brain and, eventually, to explain them. The areas shown on the brain diagram at the top and implemented in the MLC at the bottom are: Primary auditory cortex (A1), auditory belt (AB), auditory parabelt (PB), inferior prefrontal (PF), premotor (PM) and primary motor (M1) cortex. AB and PB together are sometimes called the 'auditory language area' or 'Wernicke's region' and AB and PB the 'motor language area' or 'Broca's region'.*

These simulations show that during the recognition of a word such as [pIk] the following processes take place:

1.  The auditory signal leads to stimulation of neuronal populations in superior-temporal cortex where activity spreads from primary auditory cortex, A1, to the surrounding auditory belt, AB, and parabelt, PB. This activation is mainly carried by the best-stimulated circuit(s), the target word, but partly also by its cohort members' and neighbors' circuits ([pIp], [kIk]). At the cognitive level, one may say that the system *entertains several perceptual hypotheses*.

2.  With a slight delay (realistically, 10–25 ms), activation also spreads to inferior-frontal cortex, to prefrontal, PF, premotor, PM, and (to a lesser degree) primary motor, M1, areas. This activation spreading is mainly carried by the phonological and lexical circuits best stimulated, which impose a fixed spatio-temporal pattern of activation. The advantage of such action links may lie in the separation of circuits whose perceptual parts overlap to a large degree. The syllable-initial phonemes [p] and [t] sound similar, but are based on motor programs for different articulators controlled by motor neurons at different locations in the motor system, in PM and M1, which are ~2 cm apart (Pulvermüller *et al.* 2006). Although their perceptual circuits (in A1, AB, and PB) overlap substantially, their action circuits (in PM, M1) do not to a similar degree. If circuits overlap, they cannot easily inhibit each other, a requirement for a decision and functional distinction between them. Therefore, the separation of circuit parts in the motor system enables between-circuit inhibition, and thus *facilitates a discrimination and decision process* between partly activated overlapping circuits.

3.  Activation from the most active motor circuit is fed back to superior-temporal circuit parts. The superior-temporal part of the circuit organizing the critical word [pIk] receives strong feedback activation from the action system, whereas those of competitor words receive comparably little (due to the competition process in the action system). The word-related cell assembly fully ignites; the correct word is being recognized. At the cognitive level, a *perceptual decision* has emerged.

Critically, as more and more auditory activation and information comes in, the activity hierarchy among the word-related cell assemblies shifts in favor of one; competitors are suppressed by an inhibition mechanism. Processes 1–3 involve a range of perceptual, phonological and lexical circuits, which accumulate excitation and compete simultaneously until, ultimately, activation entropy in the system decreases and one circuit ignites. Many factors, including noise, circuit overlap (cf. lexical neighborhood structure) and connection strength (cf. word frequency) — can influence the temporal dynamics of the processes. Although, the physiological word recognition process, including activation spreading, competition and ignition, may normally be very rapid (200–250 ms, see *when*-section, Hauk, Davis *et al.* 2006, Pulvermüller & Shtyrov 2006, and Pulvermüller, Shtyrov & Hauk 2009), a range of factors may lead to a delay in word recognition. Under

high entropy conditions (e.g., high noise, very strong neighbors) these processes may be delayed both in superior-temporal and inferior-frontal areas. Note, however, that the inhibition mechanism which, as argued above, is most efficient in inferior-frontal cortex, implies entropy reduction with time. Processes of generating and deciding between hypotheses analogous to the ones spelt out here in detail for speech perception are envisaged to underpin meaning comprehension and speech production as well.

The neurobiological basis of cognitive processes such as perceptual hypothesis generation and decision can be traced with explicit neurocomputational studies (Wennekers *et al.* 2006, Garagnani *et al.* 2008, Garagnani, Wennekers & Pulvermüller 2009). As mentioned, these models build strong functional links between frontal action circuits and posterior perception circuits. Therefore, a lesion in the action part of the distributed circuits does not only impair actions, it may also impact on perception and understanding, and, vice versa, lesions in the perceptual network part can reduce motor output functions in addition to causing perceptual deficits (Pulvermüller & Fadiga 2010). This is consistent, for example, with well-known reports about speech perception deficits and abnormalities in patients with different types of aphasia (Basso *et al.* 1977, Blumstein *et al.* 1994). Under taxing conditions, aphasic patients with Broca aphasia, which typically relates to frontal lesion, have difficulty understanding single words (Moineau *et al.* 2005) and even under optimal perceptual conditions, comprehension is delayed and activation of phonological cohort members reduced (Utman *et al.* 2001, Yee *et al.* 2008). Inferior-frontal lesion has a similar effect on gesture discrimination (Pazzaglia *et al.* 2008). Some colleagues chose to ignore these and similar reports of inferior-frontal lesions and their related speech perception deficits, or play them down as 'not dramatic' (Hickok 2009, Bever & Poeppel, this volume; see Pulvermüller & Fadiga 2010 for a review of this discussion). In this context, the reader should also be reminded of the TMS evidence showing that precentral stimulation alters speech perception (see *where*-section, sounds, and, for example, D'Ausilio *et al.* 2009), and of the fact that neural degeneration in the motor cortex and inferior part of the frontal lobes leads to language and conceptual deficits, especially for words and concepts related to actions (see *where*-section 2.1 and, for example, Bak *et al*. 2001).

The instant flow of activation from superior-temporal 'perceptual' neurons to inferior-frontal 'action' neurons bound together in a distributed action-perception circuit is an automatic process under standard conditions (case A: unambiguous input, good signal-to-noise ratio), but is modified under specific circumstances. In case competitor circuits are stimulated by a perceptually ambiguous stimulus ([#Ik] with # in between [p] and [t]), two lexical circuits compete and, for reaching a decision, the inferior-frontal activations and competitive mechanisms are of greatest importance. Competitor circuit activations, the associated increased entropy of the activation landscape and subsequent regulation take time so that the ignition of the winning (target) circuit will be delayed (case B). In the worst case (C), the wrong action-perception circuit may ignite first, further perceptual evidence arriving later builds up activation in the 'correct' circuit so that its ignition is much delayed and a revision of the perception process results. In all three cases, straight perception (A), high entropy perception

(B), and corrected perception (C), inferior-frontal 'action circuits' may be critical in the perception process, although the inferior-frontal activation may differ as a function of effort and attention necessary. Consistent with this view, a recent experiment on attention modulation in speech perception showed a modulatory effect of attention on inferior-frontal activation, which was stronger than the one seen in superior-temporal cortex (Shtyrov *et al.*, in press). In order to allow for similarly specific predictions, an action-perception approach — or likewise analysis-by-synthesis approach — needed to be stated in terms of a mechanistic brain-based circuit model.

Cognitive scientists entertain a major debate about the existence of symbol representations that behave in a discrete manner, becoming active in an all-or-none fashion rather than gradually. Linguistic and symbolic linguistic theories build on such discrete representations, whereas many neural processing approaches postulate distributed representations and processes that are gradual. In this case, the gradual distance of an activation pattern to the closest perceptual target vector determines the percept. Using networks fashioned according to the neuroanatomical structure and connection pattern of the left-perisylvian language cortex along with neurophysiologically realistic synaptic learning, we found neuronal correlates of discrete symbols in distributed action-perception circuits. These circuits indeed behave in a discrete fashion, showing an explosion-like ignition process to above-threshold stimulation. The circuits overlap to a degree and a specific type of realistic learning rule has a network effect of reducing this overlap (Garagnani, Wennekers & Pulvermüller 2009). The finding that distributed and discrete circuits develop in realistic neuroscience grounded networks may entail a better understanding of the mechanisms of symbol processing.

## 5.     *Why*: From Brain Mechanisms to Explanation

A range of why questions target the causal origin of brain mechanisms of language as we can infer them: Why are phonetic distinctive features mapped on specific loci of cortex? Why in the intact language-competent brain do acoustic phonological representations in superior-temporal cortex functionally depend on and interact with articulatory-phonological representations in inferior-frontal cortex? Why do phonological circuits link together into lexical ones, and why do these perisylvian cell assemblies underpinning spoken words link up with semantic networks elsewhere in the brain? Why do these semantic networks encompass so many other sets of areas, that color word connect to parahippocampal and fusiform, arm action words to precentral, and odor words to pyri-form cortex? Why do certain neuronal networks support the emergence of discrete neuronal circuits for linguistic representations? Why do some networks — dependent on their internal structure, function and learning algorithms — either support the build-up of discrete combinatorial representations with a function similar to syntactic rules or suggest that such rule-like representations do not exist? Exciting questions like these can be added almost *ad infinitum*, but it is difficult to say how far we actually are from an ultimate answer to them. Let me make an attempt to outline components of such an explanatory answer using two examples.

There are known properties of the brain and of brain function that can form a basis of neuroscientific explanation. Our brain's most important structure for cognition, the cortex, is, structurally and functionally, an associative memory (Braitenberg & Schüz 1998). Its neurons are linked by way of their synapses and links strengthen depending on their use, or correlation of activation (Tsumoto 1992). The cortex is not, however, a *tabula rasa* learning structure. It is equipped with a wealth of information. Some of this information is manifest structurally, in the anatomy of the cortex, in the structure and microstructure of areas, connections between areas and even the microstructure of neurons and their biochemical properties.

Some of the explanations and answers to 'Why X?' — questions may therefore recur to such established knowledge. As one example, explanations can be of the form 'Because X is a necessary consequence of functional correlation of neuronal processes a and b and the structural connections between the neuronal structures (neurons, neuronal assemblies) A and B'. Due to such functional correlation of structurally connected units acoustic and articulatory phonological representations connect with each other to form action-perception circuits for integration of phonological information in speech perception and production. Statistical learning of co-occurrence patterns of phonemes in words and morphemes accounts for the formation of lexical representations (Saffran *et al.* 1996, Pelucchi *et al.* 2009), which are realized as neuronal ensembles distributed over perisylvian cortex. Correlation between word form and activity in sensory and motor systems of the brain also explains the binding between sign and its referential meaning, given the relevant connections necessary for such learning are available in the first place. Information about the referents is available in different brain systems — motor, visual, auditory, olfactory, etc. — for different kinds of words, therefore the differential distribution of category-specific semantic circuits results. As soon as a stock of signs with referential meaning is available, indirect, contextual semantic learning is possible due to the correlation of new word forms with familiar ones for which referential semantic information is already available. Co-activation of the new words' neuronal circuits with semantic circuits bound to familiar words, which appear in the context of the new ones, leads to the binding of semantic neurons to the new words' circuits, thus offering a neuronal basis for in-context semantic learning (Pulvermüller 2002).

An important component of this account is correlation learning and consequent binding (i) between action and perception circuits in phonological learning, (ii) between phonological circuits in lexical learning, (iii) between word form and referential action and perception circuits in semantic learning, and (iv) between new word form circuits and previously established semantic circuits that are co-activated in contextual semantic learning. Note again that such learning is only possible if the necessary neuronal substrates and connections storing the critical correlations are available in the first place. These substrates and connections are determined to a great extent by the genetic code (Vargha-Khadem 2005, Fisher & Scharff 2009), although some influence of neuronal activity on the formation of connections cannot be denied (Rauschecker & Singer 1979, Hubel 1988, Engert & Bonhoeffer 1999, Nagerl *et al.* 2007).

For some time, modern cognitive scientists were slightly skeptical regard-

ing a major role of correlation-driven learning in cognition. This is understand-able, because, historically, modern cognitive science set itself apart from behavi-orism, which had once emphasized the role of some forms of associative learning in a tabula rasa system called a 'black box'. Evidently, such an approach is unable to explain area-specific activation patterns in the grey matter or the functional relevance of specific white matter tracts and thus cannot succeed in the neuro-biological explanation of cognitive functions. However, a dismissal of the rele-vance of correlation learning at the neurobiological level would seem an undue over-reaction to behaviorism. The empirical support for learning based on correl-ation of neuronal activity seems too strong and its implications for brain lang-uage theory too clear (see Pulvermüller 1999). Importantly, and in sharp contrast with behaviorist approaches, the effect of correlation of neuronal activation needs to be considered in the context of existing neuroanatomical connectivity.

Neuroanatomical connections of linguistic importance have been docu-mented in the healthy undeprived human brain. In particular, there are multiple connections between superior- and middle-temporal gyrus and inferior-frontal prefrontal, premotor, and opercular cortex via the capsula extrema (Saur *et al.* 2008, Petrides & Pandya 2009) and further inferior-frontal to superior-temporal connections also including inferior-prefrontal and superior-temporal areas via the fasciculus arcuatus (Glasser & Rilling 2008, Petrides & Pandya 2009). Some of these latter connections seem to stop over in the inferior-parietal cortex and in general the parietal lobe seems strongly linked into the fronto-temporal network (Catani *et al.* 2005). Some fronto-temporal connections were already evident in non-human primates (Pandya & Yeterian 1985, Petrides & Pandya 2009), although direct comparison showed that especially the arcuate fascicle is most strongly developed in humans (Rilling *et al.* 2008). This congruency and gradual difference suggests a pathway by which the genetic code did influence the emer-gence of human language in evolution. Strong direct fronto-temporal connections enabled the build-up of a large numbers of action-perception circuits for phono-logical and lexical processing in humans. As the rich fronto-temporal connections could for the first time support a great variety of fronto-temporal action-perception circuits, this proposal suggests a neurobiological explanation for the large vocabularies of human languages. Human languages include large vocabu-laries of 10,000s of spoken words (Pinker 1994), whereas our closest relatives, great apes, use only ~20–40 different signs (Pika *et al.* 2003, Tomasello & Call 2007), and even under massive training show a limit of ~200–300 symbols (Savage-Rumbaugh *et al.* 1985). As the documented fronto-temporal links seem to be mainly left-lateralized (Catani *et al.* 2005), the neuroanatomical-neuro-functi-onal approach also provides a natural explanation for why language is lateralized to the left hemisphere in most human subjects. Because the left hemisphere houses most of the fronto-temporal connections in perisylvian language cortex (Catani *et al.* 2005), action-perception correlation can best be stored there, so that the phonological and lexical action-perception circuits are lateralized to the left.

Whether aspects of syntax can also be learned in an associative manner, given some neuroanatomically manifest genetic information is available, remains a topic of debate. Investigations into statistical language learning demonstrate that much syntactic information is immanent to the correlations and conditional

probabilities between words in sentences; this information can be used, for example, to automatically classify words into lexical classes (Briscoe & Carroll 1997, Lin 1998). Neural network research demonstrates that such combinatorial information can also be extracted with neuron-like devices and neuro-functionally-inspired algorithms (Elman 1990, Honkela *et al.* 1995, Christiansen & Chater 1999, Hanson & Negishi 2002). A dispute can still occur about the degree to which the neural processes and structures can be likened to constructs postu-lated by linguists and cognitive scientists. In one view, neural networks processing syntactic information are probability mapping devices entirely dis-similar to the rule systems proposed in linguistics (Elman *et al.* 1996, McClelland & Patterson 2003).

We recently explored sequence probability mapping in neuronal networks incorporating important features of cortical connectivity frequently omitted by neural approaches (Knoblauch & Pulvermüller 2005, Pulvermüller & Knoblauch 2009). In these networks, we found formation of aggregates of neurons that, after learning, responded in a discrete 'all-or-nothing' fashion to similar contexts. These neuronal assemblies were primed by a range of past contexts and, in turn, primed a range of possible successor contexts. As an example, a range of different nouns primed the neuronal aggregate, which, in turn, activated a range of verbs. The neuronal aggregates were connected to word representations in a discrete fashion, i.e. either strongly or very weakly. Such neuronal grouping is similar to the discrete grouping of words into lexical categories (noun or verb) and the linkage of such discrete combinatorial categories bears similarity to syntactic rules linking together lexical and larger syntactic categories in a sequential fashion. A rule such as 'S → N V' or 'S → NP VP' (along with other syntactic and lexicon rules) would equally connect a wide range of utterances it covers. Inter-estingly, the combinatorial neuronal assemblies connect constituent pairs not pre-viously learned together, thus documenting a degree of generalization along with functional discreteness.

Features of the grammar network setting it apart from other models used in cognitive science to approach aspects of the serial-order problem include the following: Massive auto-associative connections within an area, neurons sensitive to sequential activation of input units, sparse activation and input coding, Hebb-type unsupervised learning, and activity control mechanisms using inhibition. These features, all of which also characterize the cortex, may contribute to the formation of combinatorial neuronal assemblies and may be important for under-standing *why* brains build rules — assuming, as some evidence suggests, that they indeed do so. The combinatorial neuronal assemblies may play a crucial role in the neuronal grammar machinery, although additional mechanisms are necessary for such a device to process a range of sentence structures (Pulvermüller 2003).

In the *why*-section of this paper, some still incomplete explanation attempts were explored, covering the laterality of language functions and its relationship to cortical connectivity, the structural and functional basis of action-perception circuits in phonological, lexical and semantic processing, and the formation of functionally discrete circuits, especially in the combinatorial domain, and the still tentative relationship of such discrete circuit emergence to network structure.

## 6.    Linguistic Summary and Synopsis

From the standpoint of linguistic theory, what is gained from recent neuroscience research? In the semantic domain, we have learned that there is a very real sense in which semantic categories exist. Meaning is not the mental representation of objects, relevant to it are action aspects as well. At the semantic level, language is 'woven into action' and this insight from the analytical theory of language is backed by brain research. Motor and sensory systems activation demonstrates semantic categories along brain dimensions. Additional areas, in the vicinity of sensorimotor domains, may play a role in abstract semantic processing and in general meaning access.

In a similar vein, phonological distinctions can be objectified based on brain correlates. Phonetic distinctive features have their correlates in local cortical activation in the auditory and motor systems. This addresses questions about the nature of phonological representation: Should phoneme features be construed as articulatory or as acoustic? In brain terms, they are both as the phonological circuits appear to link motor and auditory circuits with each other.

An intensive debate about the nature of mental computation can be addressed based on the results from the neuroscience of language. Neuronal ensemble theory along with empirical neurophysiological evidence supports the existence of discrete cortical representations and mechanistic underpinnings for rules of grammar. The position once backed by neural network simulations that rules do not exist at the neuronal level may be in need of revision.

The idea that it takes about half a second to understand a word or sentence — counted from the point in time when the last word critical for sentence understanding is first unambiguously present in the input — might imply a substantial delay in the comprehension process and, as discussed above, one may wonder whether such a delay could represent a substantial disadvantage biologically. Supportive of rapid, almost instantaneous understanding comes from recent neurophysiological studies suggesting latencies of <250 ms of the earliest brain correlates of semantic word and sentence understanding and syntactic parsing. These neurophysiological results support rapid and parallel psycholinguistic models and argue against slow-serial or -cascaded theories assuming sequential steps from phonological to syntactic and semantic modules of hundreds of milliseconds. Relevant time delays seem to range around 10–50 ms only, thus indicating near-simultaneous activation and information access.

Looking back at the review, progress in the *where-* and *when*-domain is certainly most impressive. In my view, however, the maturity of the field, its stage of development, will be evaluated in light of plausible approaches to *how* and *why* issues. Collecting wisdom about new plants, stars, and brain activation loci can advance a field in a hunter-gatherer sense. In order for it to transform into an explanatory science, *explanations* need to be offered (Hempel & Oppenheim 1948, von Wright 1971). In the neuroscience of language, these explanations use neuroscience facts and established principles of brain structure and function as *explanans*. It is in this explanatory domain where, in my view, further progress is most desperately needed. Some little progress has been made, which, however, lacks the flashy aspect of newly discovered neurocognitive hotspots. An im-

portant achievement, now and in the future, may therefore be neuromechanistic explanations detailing *why* specific brain areas are necessary for, or light up and index, specific facets of language processing, *how* neuronal ensembles and distributed areas become activated with precisely timed milli-second delays, and which precise neuronal wirings can potentially account for neurometabolic activation of specific cortical clusters in semantic understanding.

## References

Aho, Alfred V., John E. Hopcroft & Jeffrey D. Ullman. 1987. *Data Structure and Algorithms*. Reading, MA: Addison-Wesley.

Araujo, Ivan E. de, Edmund T. Rolls, Maria Inés Velazco, Christian Margot & Isabelle Cayeux. 2005. Cognitive modulation of olfactory processing. *Neuron* 46, 671–679.

Assadollahi, Ramin & Friedemann Pulvermüller. 2001. Neuromagnetic evidence for early access to cognitive representations. *Neuroreport* 12, 207–213.

Aziz-Zadeh, Lisa, Stephen M. Wilson, Giacomo Rizzolatti & Marco Iacoboni. 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology* 16, 1818–1823.

Bak, Thomas H., Despina Yancopoulou, Peter J. Nestor, John H. Xuereb, Maria G. Spillantini, Friedemann Pulvermüller & John R. Hodges. 2006. Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain* 129 (pt. 2), 321–332.

Bak, Thomas H., Dominic G. O'Donovan, John H. Xuereb, Simon Boniface & John R. Hodges. 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the Motor Neurone Disease–Dementia–Aphasia syndrome. *Brain* 124, 103–120.

Barber, Horacio A. & Marta Kutas. 2007. Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Review* 53, 98–123.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22, 577–609; discussion 610–560.

Basso, Anna, G. Casati & L. A. Vignolo. 1977. Phonemic identification defect in aphasia. *Cortex* 13, 85–95.

Bedny, Marina, Alfonso Caramazza, Emily Grossman, Alvaro Pascual-Leone & Rebecca Saxe. 2008. Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience* 28, 11347–11353.

Berkum, Jos J. van, Colin M. Brown, Peter Hagoort & Pienie Zwitserlood. 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology* 40, 235–248.

Berkum, Jos J. van, Pienie Zwitserlood, Peter Hagoort & Colin M. Brown. 2003. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research* 17, 701–718.

Bever, Thomas G. & David Poeppel. This volume. Analysis by synthesis: A (re-)

emerging program of research for language and vision.

Binder, Jeffrey R., Chris F. Westbury, K. A. McKiernan, Edward T. Possing & David A. Medler. 2005. Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience* 17, 905–917.

Blumstein, Sheila E., Martha Burton, Shari Baum, Robin Waldstein & Donald Katz. 1994. The role of lexical status on the phonetic categorization of speech in aphasia. *Brain and Language* 46, 181–197.

Bornkessel, Ina, Matthias Schlesewsky & Angela D. Friederici. 2002. Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition* 85, B21–30.

Boulenger, Véronique, Olaf Hauk & Friedemann Pulvermüller. 2009. Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex* 19, 1905–1914.

Braitenberg, Valentino & Almut Schüz. 1998. *Cortex: Statistics and Geometry of Neuronal Connectivity*, 2 edn. Berlin: Springer.

Briscoe, Edward & John Carroll. 1997. Automatic extraction of subcategorization from corpora. *Proceedings of the 5th ACL Conference on Applied Natural Language Processing, Washington, DC, 5*, 1–9.

Brown, Colin M., Jos J. van Berkum & Peter Hagoort. 2000. Discourse before gender: An event-related brain potential study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholinguistic Research* 29, 53–68.

Bruce, Charles J. & Michael E. Goldberg. 1985. Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology* 53, 603–635.

Caplan, David, Nathaniel Alpert, Gloria Waters & Anthony Olivieri. 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping* 9, 65–71.

Caplan, David, Louise Stanczak & Gloria Waters. 2008. Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. *Journal of Cognitive Neuroscience* 20, 643–656.

Catani, Marco, Derek K. Jones & Dominic H. Ffytche. 2005. Perisylvian language networks of the human brain. *Annals of Neurology* 57, 8–16.

Chao, Linda L., James V. Haxby & Alex Martin. 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience* 2, 913–919.

Christiansen, Morten H. & Nick Chater. 1999. Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23, 157–205.

D'Ausilio, Alessandro, Friedemann Pulvermüller, Paola Salmas, Ilaria Bufalari, Chiara Begliomini & Luciano Fadiga. 2009. The motor somatotopy of speech perception. *Current Biology* 19, 381–385.

Damasio, Antonio R. & Daniel Tranel. 1993. Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences, USA, 90*, 4957–4960.

Damasio, Hanna, Thomas J. Grabowski, Daniel Tranel, Richard D. Hichwa & Antonio R. Damasio. 1996. A neural basis for lexical retrieval. *Nature* 380, 499–505.

Daniele, Antonio, Laura Giustolisi, M. Caterina Silveri, Cesare Colosimo & Guido Gainotti. 1994. Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. *Neuropsychologia* 32, 1325–1341.

Dehaene-Lambertz, Ghislaine. 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919–924.

Diesch, Eugen, Carsten Eulitz, Scott Hampson & Bernhard Ross. 1996. The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain and Language* 53, 143–168.

Elman, Jeffrey L. 1990. Finding structure in time. *Cognitive Science* 14, 179–211.

Elman, Jeffrey L., Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Engert, Florian & Tobias Bonhoeffer. 1999. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature* 399, 66–70.

Epstein, Herman T. 1999. Other brain effects of words. *Behavioral and Brain Sciences* 22, 287–288.

Fadiga, Luciano, Laila Craighero, Giovanni Buccino & Giacomo Rizzolatti. 2002. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15, 399–402.

Fadiga, Luciano, Leonardo Fogassi, Giovanni Pavesi & Giacomo Rizzolatti. 1995. Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73, 2608–2611.

Fisher, Simon E. & Constance Scharff. 2009. FOXP2 as a molecular window into speech and language. *Trends in Genetics* 25, 166–177.

Fodor, Jerry A. 1983. *The Modulatity of Mind*. Cambridge, MA: MIT Press.

Friederici, Angela D. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6, 78–84.

Friederici, Angela D. 2009. Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences* 13, 175–181.

Friederici, Angela D., Martin Meyer & D. Yves von Cramon. 2000. Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language* 74, 289–300.

Friederici, Angela D., Erdmut Pfeifer & Anja Hahne. 1993. Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1, 183–192.

Gainotti, Guido. 2006. Anatomical functional and cognitive determinants of semantic memory disorders. *Neuroscience and Biobehavioral Reviews* 30, 577–594.

Gainotti, Guido. 2008. Chapter 9 Disorders of semantic memory. *Handbook of Clinical Neurology* 88, 203–223.

Garagnani, Max, Yury Shtyrov & Friedemann Pulvermüller. 2009. Effects of attention on what is known and what is not: MEG evidence for functionally discrete memory circuits. *Frontiers in Human Neuroscience* 3, doi:10.3389/neuro.3309.3010.2009.

Garagnani, Max, Thomas Wennekers & Friedemann Pulvermüller. 2007. A neuronal model of the language cortex. *Neurocomputing* 70, 1914–1919.

Garagnani, Max, Thomas Wennekers & Friedemann Pulvermüller. 2008. A neuro-

anatomically-grounded Hebbian learning model of attention-language interactions in the human brain. *European Journal of Neuroscience* 27, 492–513.

Garagnani, Max, Thomas Wennekers & Friedemann Pulvermüller. 2009. Recruitment and consolidation of cell assemblies for words by way of Hebbian learning and competition in a multi-layer neural network. *Cognitive Computation* 1, 160–176.

Gaskell, M. Gareth & William D. Marslen-Wilson. 2002. Representation and competition in the perception of spoken words. *Cognitive Psychology* 45, 220–266.

Geschwind, Norman. 1970. The organization of language and the brain. *Science* 170, 940–944.

Glasser, Matthew F. & James K. Rilling. 2008. DTI tractography of the human brain's language pathways. *Cerebral Cortex* 18, 2471–2482.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7, 219–224.

González, Julio, Alfonso Barros-Loscertales, Friedemann Pulvermüller, Vanessa Meseguer, Ana Sanjuán, Vicente Belloch & César Ávila. 2006. Reading 'cinnamon' activates olfactory brain regions. *Neuroimage* 32, 906–912.

Griffiths, Timothy D. & Jason D. Warren. 2002. The planum temporale as a computational hub. *Trends in Neurosciences* 25, 348–353.

Halle, Morris & Kenneth N. Stevens. 1959. Analysis by synthesis. In W. Wathen-Dunn & L. E. Woods (eds.), *Proceedings of the Seminar on Speech Compression and Processing*, vol. 2, Section D7.

Halle, Morris & Kenneth N. Stevens. 1962. Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8, 155–159.

Hanson, Stephen José & Michiro Negishi. 2002. On the emergence of rules in neural networks. *Neural Computation* 14, 2245–2268.

Hare, Mary, Jeffrey L. Elman & Kim G. Daugherty. 1995. Default generalisation in connectionist networks. *Language and Cognitive Processes* 10, 601–630.

Hasting, Anna S. & Sonja A. Kotz. 2008. Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience* 20, 1207–1219.

Hauk, Olaf & Friedemann Pulvermüller. 2004. Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping* 21, 191–201.

Hauk, Olaf, Ingrid Johnsrude & Friedemann Pulvermüller. 2004. Somatotopic representation of action words in the motor and premotor cortex. *Neuron* 41, 301–307.

Hauk, Olaf, Matthew H. Davis, Ferath Kherif & Friedemann Pulvermüller. 2008. Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *European Journal of Neuroscience* 27, 1856–1866.

Hauk, Olaf, Matthew H. Davis, M. Ford, Friedemann Pulvermüller & William D. Marslen-Wilson. 2006. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage* 30, 1383–1400.

Hauk, Olaf, Yury Shtyrov & Friedemann Pulvermüller. 2006. The sound of

actions as reflected by mismatch negativity: Rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *European Journal of Neuroscience* 23, 811–821.

Hempel, Carl Gustav & Paul Oppenheim. 1948. Studies in the logic of explanation. *Philosophy of science* 15, 135–175.

Hesselmann, Volker, Bettina Sorger, Kathrin Lasek, Orlando Guntinas-Lichius, Barbara Krug, Volker Sturm, Rainer Goebel & Klaus Lackner. 2004. Discriminating the cortical representation sites of tongue and up movement by functional MRI. *Brain Topography* 16, 159–167.

Hickok, Gregory. 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21, 1229–1243.

Hickok, Gregory & David Poeppel. 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402.

Hodges, John R & Karalyn Patterson. 2007. Semantic dementia: A unique clinico-pathological syndrome. *Lancet Neurology* 6, 1004–1014.

Honkela, Timo, Ville Pulkki & Teuvo Kohonen. 1995. Contextual relations of words in Grimm tales analyzed by self-organizing map. *Proceedings of the International Conference on Neural Networks (ICANN) 1995*, 3–7.

Hubel, David. 1988. *Eye, Brain, and Vision*. New York: Freeman.

Indefrey, Peter, Colin M. Brown, Frauke Hellwig, Katrin Amunts, Hans Herzog, Rüdiger J. Seitz & Peter Hagoort. 2001. A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences, USA*, 98, 5933–5936.

Just, Marcel Adam, Patricia A. Carpenter, Timothy A. Keller, William F. Eddy & Keith R. Thulborn. 1996. Brain activation modulated by sentence comprehension. *Science* 274, 114–116.

Kemmerer, David & Javier Gonzalez-Castillo. 2010. The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language* 112, 54–76.

Kempen, Gerard & Karin Harbusch. 2003. An artificial opposition between grammaticality and frequency: Comment on Bornkessel, Schlesewsky & Friederici (2002). *Cognition* 90, 205–210; discussion 211–203, 215.

Kiefer, Markus, Eun-Jin Sim, Bärbel Herrnberger, Jo Grothe & Klaus Hoenig. 2008. The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *The Journal of Neuroscience* 28, 12224–12230.

Kiefer, Markus, Eun-Jin Sim, Sarah Liebich, Olaf Hauk & James Tanaka. 2007. Experience-dependent plasticity of conceptual representations in human sensory-motor areas. *Journal of Cognitive Neuroscience* 19, 525–542.

Knoblauch, Andreas & Friedemann Pulvermüller. 2005. Sequence detector networks and associative learning of grammatical categories. In Stefan Wermter, Günther Palm & Mark Elshaw (eds.), *Biomimetic Neural Learning for Intelligent Robots*, 31–53. Berlin: Springer.

Lahav, Amir, Elliot Saltzman & Gottfried Schlaug. 2007. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience* 27, 308–314.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal*

*about the Mind*. Chicago, IL: University of Chicago Press.

Landau, Barbara & Lila R. Gleitman. 1985. *Language and Experience: Evidence from the Blind Child*. Cambridge, MA: Harvard University Press.

Landauer, Thomas K. & Susan T. Dumais. 1997. A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104, 211–240.

Lee, Bernard S. 1950. Effects of delayed speech feedback. *The Journal of the Acoustical Society of America* 22, 824–826.

Lin, Dekang. 1998. Automatic retrieval and clustering of similar words. *Proceedings of the 17th International Conference on Computational Linguistics & 36th Annual Meeting of the Association for Computational Linguistics,* 768–773.

Lotto, Andrew J., Gregory S. Hickok & Lori L. Holt. 2009. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences* 13, 110–114.

Lotze, Martin, G. Seggewies, Michael Erb, Wolfgang Grodd & Niels Birbaumer. 2000. The representation of articulation in the primary sensorimotor cortex. *Neuroreport* 11, 2985–2989.

Mahon, Bradford Z., Stefano Anzellotti, Jens Schwarzbach, Massimiliano Zampini & Alfonso Caramazza. 2009. Category-specific organization in the human brain does not require visual experience. *Neuron* 63, 397–405.

Mahon, Bradford Z. & Alfonso Caramazza. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology Paris* 102, 59–70.

Marslen-Wilson, William D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25, 71–102.

Martin, Alex. 2007. The representation of object concepts in the brain. *Annual Review of Psychology* 58, 25–45.

Martin, Alex, James V. Haxby, François M. Lalonde, Cheri L. Wiggs & Leslie G. Ungerleider. 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270, 102–105.

McClelland, James L. & Karalyn Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6, 465–472.

McClelland, James L. & Karalyn Patterson. 2003. Differentiation and integration in human language. Reply to Marslen-Wilson & Tyler. *Trends in Cognitive Sciences* 7, 63–64.

Miceli, Gabriele, Erin Fouch, Rita Capasso, Jennifer R. Shelton, Francesco Tomaiuolo & Alfonso Caramazza. 2001. The dissociation of color from form and function knowledge. *Nature Neuroscience* 4, 662–667.

Moineau, Suzanne, Nina F. Dronkers & Elizabeth Bates. 2005. Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language and Hearing Research* 48, 884–896.

Moscoso Del Prado Martin, Fermin, Olaf Hauk & Friedemann Pulvermüller. 2006. Category specificity in the processing of color-related and form-related words: An ERP study. *Neuroimage* 29, 29–37.

Näätänen, Risto, Anne Lehtokoski, Mietta Lennes, Marie Cheour, Minna Huotilainen, Antti Iivonen, Martti Vainio, Paavo Alku, Risto J. Ilmoniemi, Aavo Luuk, Jüri Allik, Janne Sinkkonen & Kimmo Alho. 1997. Language-specific

phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.

Nagerl, U. Valentin, German Köstinger, John C. Anderson, Kevan A. Martin & Tobias Bonhoeffer. 2007. Protracted synaptogenesis after activity-dependent spinogenesis in hippocampal neurons. *The Journal of Neuroscience* 27, 8149–8156.

Neininger, Bettina & Friedemann Pulvermüller. 2003. Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia* 41, 53–70.

Neville, Helen, Janet L. Nicol, Andrew Barss, Kenneth I. Forster & Merrill F. Garrett. 1991. Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3, 151–165.

Noppeney, Uta & Cathy J. Price. 2004. Retrieval of abstract semantics. *Neuroimage* 22, 164–170.

Obleser, Jonas, Henning Boecker, Alexander Drzezga, Bernhard Haslinger, Andreas Hennenlotter, Michael Roettinger, Carsten Eulitz & Josef P. Rauschecker. 2006. Vowel sound extraction in anterior superior temporal cortex. *Human Brain Mapping* 27, 562–571.

Obleser, Jonas, Aditi Lahiri & Carsten Eulitz. 2003. Auditory-evoked magnetic field codes place of articulation in timing and topography around 100 milliseconds post syllable onset. *Neuroimage* 20, 1839–1847.

Obleser, Jonas, Jonas Zimmermann, John Van Meter & Josef P. Rauschecker. 2007. Multiple stages of auditory speech perception reflected in event-related FMRI. *Cerebral Cortex* 17, 2251–2257.

Oliveri, Massimiliano, Chiara Finocchiaro, Kevin Shapiro, Massimo Gangitano, Alfonso Caramazza & Alvaro Pascual-Leone. 2004. All talk and no action: A transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience* 16, 374–381.

Pandya, Deepak N. & E. H. Yeterian. 1985. Architecture and connections of cortical association areas. In A. Peters & E. G. Jones (eds.), *Cerebral Cortex*, vol. 4: *Association and Auditory Cortices*, 3–61. London: Plenum Press.

Patterson, Karalyn, Peter J. Nestor & Timothy T. Rogers. 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience* 8, 976–987.

Paus, Tomas, David W. Perry, Robert J. Zatorre, Keith J. Worsley & Alan C. Evans. 1996. Modulation of cerebral blood flow in the human auditory cortex during speech: Role of motor-to-sensory discharges. *European Journal of Neuroscience* 8, 2236–2246.

Pazzaglia, Mariella, Nicola Smania, Elisabetta Corato & Salvatore Maria Aglioti. 2008. Neural underpinnings of gesture discrimination in patients with limb apraxia. *Journal of Neuroscience* 28, 3030–3041.

Pelucchi, Bruna, Jessica F. Hay & Jenny R. Saffran. 2009. Statistical learning in a natural language by 8-month-old infants. *Child Development* 80, 674–685.

Penolazzi, Barbara, Olaf Hauk & Friedemann Pulvermüller. 2007. Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology* 74, 374–388.

Petrides, Micahel & Deepak N. Pandya. 2009. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biology* 7,

e1000170.

Pika, Simone, Katja Liebal & Michael Tomasello. 2003. Gestural communication in young gorillas (Gorilla gorilla): Gestural repertoire, learning, and use. *American Journal of Primatology* 60, 95–111.

Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language.* New York: Harper Collins Publishers.

Pinker, Steven & Michael T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6, 456–463.

Posner, Michael I. & Gregory J. DiGirolamo. 1999. Flexible neural circuitry in word processing. *Behavioral and Brain Sciences* 22, 299–300.

Postle, Natasha, Katie L. McMahon, Roderick Ashton, Matthew Meredith & Greig I. de Zubicaray. 2008. Action word meaning representations in cyto-architectonically defined primary and premotor cortices. *Neuroimage* 43, 634–644.

Praamstra, Peter & Dick F. Stegeman. 1993. Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research* 1, 73–86.

Pulvermüller, Friedemann. 1999. Words in the brain's language. *Behavioral and Brain Sciences* 22, 253–336.

Pulvermüller, Friedemann. 2002. A brain perspective on language mechanisms: From discrete neuronal ensembles to serial order. *Progress in Neurobiology* 67, 85–111.

Pulvermüller, Friedemann. 2003. *The Neuroscience of Language.* Cambridge: Cambridge University Press.

Pulvermüller, Fridemann. 2005. Brain mechanisms linking language and action. *Nature Reviews Neuroscience* 6, 576–582.

Pulvermüller, Friedemann. 2008. Grounding language in the brain. In Manuel de Vega, Arthur Graesser & Arthur M. Glenberg (eds.), *Symbols, Embodiment, and Meanin*, 85–116. Oxford: Oxford University Press.

Pulvermüller, Friedemann. 2010. Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain and Language* 112, 167–179.

Pulvermüller, Friedemann & Andreas Knoblauch. 2009. Discrete combinatorial circuits emerging in neural networks: A mechanism for rules of grammar in the human brain? *Neural Networks* 22, 161–172.

Pulvermüller, Friedemann & Bettina Mohr. 1996. The concept of transcortical cell assemblies: A key to the understanding of cortical lateralization and inter-hemispheric interaction. *Neuroscience and Biobehavioral Reviews* 20, 557–566.

Pulvermüller, Friedemann & Luciano Fadiga. 2010. Active perception: Sensori-motor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11, 351–360.

Pulvermüller, Friedemann & Olaf Hauk. 2006. Category-specific processing of color and form words in left fronto-temporal cortex. *Cerebral Cortex* 16, 1193–1201.

Pulvermüller, Friedemann & Ramin Assadollahi. 2007. Grammar or serial order? Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. *Journal of Cognitive Neuroscience* 19, 971–980.

Pulvermüller, Friedemann & Yury Shtyrov. 2006. Language outside the focus of

attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology* 79, 49–71.

Pulvermüller, Friedemann & Yury Shtyrov. 2009. Spatio-temporal signatures of large-scale synfire chains for speech: MEG evidence. *Cerebral Cortex* 19, 79–88.

Pulvermüller, Friedemann, Elisa Cooper-Pye, Clare Dine, Olaf Hauk, Peter J. Nestor & Karalyn Patterson. In press. The word processing deficit in Semantic Dementia: All categories are equal but some categories are more equal than others. *Journal of Cognitive Neuroscience*.

Pulvermüller, Friedemann, Ferath Kherif, Olaf Hauk, Bettina Mohr & Ian Nimmo-Smith. 2009. Cortical cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Human Brain Mapping* 30, 3837–3850.

Pulvermüller, Friedemann, Markus Härle & Friedhelm Hummel. 2000. Neuro-physiological distinction of verb categories. *Neuroreport* 11, 2789–2793.

Pulvermüller, Friedemann, Martina Huss, Ferath Kherif, Fermin Moscoso del Prado Martin, Olaf Hauk & Yury Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences, USA*, 103, 7865–7870.

Pulvermüller, Friedemann, Olaf Hauk, Vadim V. Nikulin & Risto J. Ilmoniemi. 2005. Functional links between motor and language systems. *European Journal of Neuroscience* 21, 793–797.

Pulvermüller, Friedemann, Werner Lutzenberger & Hubert Preissl. 1999. Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex* 9, 498–508.

Pulvermüller, Friedemann, Yury Shtyrov & Olaf Hauk. 2009. Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language* 110, 81–94.

Pulvermüller, Friedemann, Yury Shtyrov & Risto J. Ilmoniemi. 2003. Spatio-temporal patterns of neural language processing: An MEG study using Minimum-Norm Current Estimates. *Neuroimage* 20, 1020–1025.

Pulvermüller, Friedemann, Yury Shtyrov & Risto J. Ilmoniemi. 2005. Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience* 17, 884–892.

Pulvermüller, Friedemann, Yury Shtyrov, Anna Hasting & Robert P. Carlyon. 2008. Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language* 104, 244–253.

Raposo, Ana, Helen E. Moss, Emmanuel A. Stamatakis & Lorraine K. Tyler. 2009. Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47, 388–396.

Rauschecker, Josef P. & Sophie K. Scott. 2009. Maps and streams in the auditory cortex: Non-human primates illuminate human speech processing. *Nature Neuroscience* 12, 718–724.

Rauschecker, Josef P. & Wolf Singer. 1979. Changes in the circuitry of the kitten visual cortex are gated by postsynaptic activity. *Nature* 280, 58–60.

Rilling, James K., Matthew F. Glasser, Todd M. Preuss, Xiangyang Ma, Tiejun Zhao, Xiaoping Hu & Timothy E. J. Behrens. 2008. The evolution of the

arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience* 11, 426–428.

Rizzolatti, Giacomo & Corrado Sinigaglia. 2010. The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience* 11, 264–274.

Rizzolatti, Giacomo & Laila Craighero. 2004. The mirror-neuron system. *Annual Review in Neuroscience* 27, 169–192.

Romanski, Lizabeth M., Biao Tian, Jonathan B. Fritz, Mortimer Mishkin, Patricia S. Goldman-Rakic & Josef P. Rauschecker. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience* 2, 1131–1136.

Rugg, Michael D. 1984. Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia* 22, 435–443.

Rumelhart, David E. & James L. McClelland. 1987. Learning the past tense of English verbs: Implicit rules or parallel distributed processing. In Brian MacWhinney (ed.), *Mechanisms of Language Acquisition*, 194–248. Hillsdale, NJ: Erlbaum.

Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Salmelin, Riitta, Päivi Helenius & Kari Kuukka. 1999. Only time can tell — words in context. *Behavioral and Brain Sciences* 22, 300.

Saur, Dorothee, Björn W. Kreher, Susanne Schnell, Dorothee Kümmerer, Philipp Kellmeyer, Magnus-Sebastian Vry, Roza Umarova, Mariacristina Musso, Volkmar Glauche, Stefanie Abeld, Walter Huberd, Michel Rijntjes, Jürgen Hennig & Cornelius Weillera. 2008. Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences, USA* 105, 18035–18040.

Savage-Rumbaugh, Sue, Duane M. Rumbaugh & Kelly McDonald. 1985. Language learning in two species of apes. *Neuroscience Biobehavioral Reviews* 9, 653–665.

Scott, Sophie K. & Ingrid S. Johnsrude. 2003. The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences* 26, 100–107.

Scott, Sophie K., C. Catrin Blank, Stuart Rosen & Richard J. Wise. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, Pt 12, 2400–2406.

Scott, Sophie K., Carolyn McGettigan & Frank Eisner. 2009. A little more conversation, a little less action — Candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience* 10, 295–302.

Shtyrov, Yury, Friedemann Pulvermüller, Risto Näätänen & Risto J. Ilmoniemi. 2003. Grammar processing outside the focus of attention: An MEG study. *Journal of Cognitive Neuroscience* 15, 1195–1206.

Shtyrov, Yury, Olaf Hauk & Friedemann Pulvermüller. 2004. Distributed neuronal networks for encoding category-specific semantic information: The mismatch negativity to action words. *European Journal of Neuroscience* 19, 1083–1092.

Shtyrov, Yury, Teija Kujala & Friedemann Pulvermüller. In press. Interactions between language and attention systems: Early automatic lexical processing? *Journal of Cognitive Neuroscience*.

Sim, Eun-Jin & Markus Kiefer. 2005. Category-related brain activity to natural categories is associated with the retrieval of visual features: Evidence from repetition effects during visual and functional judgments. *Cognitive Brain Research* 24, 260–273.

Simmons, W. Kyle, Vimal Ramjee, Michael S. Beauchamp, Ken McRae, Alex Martin & Lawrence W. Barsalou. 2007. A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45, 2802–2810.

Skrandies, Wolfgang. 1999. Early effects of semantic meaning on electrical brain activity. *Behavioral and Brain Sciences* 22, 301.

Straube, Thomas, Hans-Joachim Mentzel, Madlen Glauer & Wolfgang H. Miltner. 2004. Brain activation to phobia-related words in phobic subjects. *Neuroscience Letters* 372, 204–208.

Tettamanti, Marco, Giovanni Buccino, Maria Cristina Saccuman, Vittorio Gallese, Massimo Danna, Paola Scifo, Ferruccio Fazio, Giacomo Rizzolatti, Stefano F. Cappa & Daniela Perani. 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17, 273–281.

Tomasello, Michael & Josep Call. 2007. Ape gestures and the origins of language. In Josep Call & Michael Tomasello (eds.), *The Gestural Communication of Apes and Monkeys*, 221–239. Mahwah, NJ: Lawrence Erlbaum Associates.

Tomasino, Barbara, Cornelius J. Werner, Peter H. Weiss & Gereon R. Fink. 2007. Stimulus properties matter more than perspective: an fMRI study of mental imagery and silent reading of action phrases. *Neuroimage* 36, Suppl 2, T128–141.

Tranel, Daniel & Antonio R. Damasio. 1999. The neurobiology of knowledge retrieval. *Behavioral and Brain Sciences* 22, 303.

Tranel, Daniel, David Kemmerer, Ralph Adolphs, Hanna Damasio & Antonio R. Damasio. 2003. Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology* 20, 409–432.

Tsumoto, Tadaharu. 1992. Long-term potentiation and long-term depression in the neocortex. *Progress in Neurobiology* 39, 209–228.

Uppenkamp, Stefan, Ingrid S. Johnsrude, Dennis Norris, William Marslen-Wilson & Roy D. Patterson. 2006. Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* 31, 1284–1296.

Utman, Jennifer Aydelott, Sheila E. Blumstein & Kelly Sullivan. 2001. Mapping from sound to meaning: Reduced lexical activation in Broca's aphasics. *Brain and Language* 79, 444–472.

Vargha-Khadem, Faraneh, David G. Gadian, Andrew Copp & Mortimer Mishkin. 2005. FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience* 6, 131–138.

Warren, Paul & William Marslen-Wilson. 1987. Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics* 41, 262–275.

Warren, Paul & William Marslen-Wilson. 1988. Cues to lexical choice: Discriminating place and voice. *Perception and Psychophysics* 43, 21–30.

Warrington, Elizabeth K. & Rosaleen A. McCarthy. 1983. Category specific access dysphasia. *Brain* 106, 859–878.

Warrington, Elizabeth K. & Tim Shallice. 1984. Category specific semantic impair-

ments. *Brain* 107, 829–854.

Wennekers, Thomas, Max Garagnani & Friedemann Pulvermüller. 2006. Language models based on Hebbian cell assemblies. *Journal of Physiology Paris* 100, 16–30.

Wilson, Stephen M., Ayşe P. Saygin, Marty I. Sereno & Marco Iacoboni. 2004. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience* 7, 701–702.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Wright, Georg H. von 1971. *Explanation and Understanding*. Ithaca, NY: Cornell University Press.

Yee, Eiling, Sheila E. Blumstein & Julie C. Sedivy. 2008. Lexical-semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *Journal of Cognitive Neuroscience* 20, 592–612.

*Friedemann Pulvermüller*
*MRC Cognition and Brain Sciences Unit*
*15 Chaucer Road*
*Cambridge CB2 7EF*
*United Kingdom*
*friedemann.pulvermuller@mrc-cbu.cam.ac.uk*

# Why Possibly Language Evolved

Peter J. Richerson  &  Robert Boyd

Human language has no close parallels in other systems of animal communication. Yet it is an important part of the cultural adaptation that serves to make humans an exceedingly successful species. In the past 20 years, a diverse set of evolutionary scholars have tried to answer the question of how language evolved in our species and why it is unique to us. They have converged on the idea that the cultural and innate aspects of language were tightly linked in a process of gene-culture coevolution. They differ widely about the details of the process, particularly over the division of labor between genes and culture in the coevolutionary process. Why is language restricted to humans given that communication seems to be so useful? A plausible answer is that language is part of human cooperation. Why did the coevolutionary process come to rest leaving impressive cultural diversity in human languages? A plausible answer is that language diversity functions to limit communication between people who cannot freely trust one another or where even truthful communications from others would result in maladaptive behavior on the part of listeners.

## 1.    Introduction

Humans are highly unusual animals in depending upon social transmission from others for acquiring most of their adaptations (Boyd & Richerson 2005, Richerson & Boyd 2005), including the specific languages they speak. Language is essential to our complex social life, revolving as it does around institutions that are transmitted by language and operated by oratory. The evolution of language as a human capacity, and of languages themselves, are subjects of a large and growing literature. Many important and controversial issues are under examination including the roles of cultural and genetic evolution in the process and the role of general purpose versus language specific innate cognitive resources. Culture itself evolves by processes that are something like the evolution of genes, but which are different in many important details. In the case of language, the cultural evolutionary processes by which

languages change over time are tolerably well understood from the work of socio-linguists (Labov 1994, Labov 2001) and historical linguists (Deutscher 2005). The application of formal phylogenetic methods borrowed from evolutionary biology to the reconstruction of language evolution is an active area of research (Gray *et al.* 2009, Pagel 2009). We have reasonable general models of cultural evolution (Henrich & McElreath 2008). We have good models of the coevolutionary process (Richerson & Boyd 1989, Feldman & Laland 1996).

The coevolutionary approach to the evolution of language was first articulated by Pinker & Bloom (1990). They imagine that the first step in the evolution of human language would be a rudimentary culturally transmitted set of signals. If having such signals were adaptively advantageous, selection might fall in genes to expand the capacity to acquire such signals. So long as a higher capacity cultural communi-cation system was favored, cognitive modifications to more efficiently acquire expli-citly linguistic features like symbolic words and grammar would be favored as the system passed some threshold of complexity. Eventually, languages with a rather large vocabulary and complex syntax and/or morphology became cognitively possible. Coevolution assumes that some relatively easy and gradual path was available such that the evolution of language could proceed from simple vocal and gestural communications to human language by some combination of small cultural and innate steps, at least after some key cognitive precursors had evolved (Origgi & Sperber 2000). See Donald (1991), Deacon (1997), Tomasello (2008), and Progovac (this volume) on how language might have evolved gradually.

Pinker (2003) continues to argue that humans were equipped by the coevolutionary process with specialized innate mechanisms to manage language acquisition, production, and comprehension. Others have argued that more general cognitive resources, perhaps especially resources shared by other domains of culture, can underpin language. We have formal models of how language specific features like compositionality might evolve by cultural evolution by languages adapting to be learnable using general cognitive resources (de Boer 2000, Munroe & Cangelosi 2002, Smith & Kirby 2008). Steels (2009) has simulated how language might be invented using sufficiently powerful general cognitive resources using laboratory robots. Briscoe (2009) argues that even if general cognitive resources are sufficient to initiate simple languages, coevolution would have produced at least some language specific biases and constraints to make langue more efficiently learnable. Of course, some genetic changes must have accompanied the evolution of language since even chimpanzees and bonobos when raised in a linguistic environment develop, at best, a rudimentary form of language (Savage-Rumbaugh & Lewin 1994). Tomasello (2008) emphasizes the idea of a cognitive complex of shared attention, collaborative activities, social motivation and cultural evolution that is important for language *and* technical and social skills. Bloom (2000) emphasized the cognitive strategies used to learn words. Dehaene's (2009) interesting work on reading shows how visual circuits evolved to process normal visual stimuli can be 'recycled' to transmit linguistic information in the visual mode, providing a possible

model for earlier coevolutionary events. Nettle (2007) raises the possibility that this sort of coevolution still goes on. Subtle differences in innate capacities may have coevolved with different language families; for example, populations with a long history of using tone to make linguistic distinctions may find it slightly easier to learn and use tones than populations not so exposed (Dediu & Ladd 2007). Good evidence to this effect would suggest that the coevolutionary process indeed affects language evolution. Thus, while a consensus exists that human language evolved by a coevolutionary scenario, evolutionary linguists differ greatly on the details.

This bare-bones account leaves many unanswered questions. Two seem fundamental to us. First, why is human language unique? Not only primates but also parrots and songbirds have preadaptations for vocal communication and vocal imitation (Baptista & Trail 1992, Pepperberg 2000). Human language is widely counted as our most spectacular adaptation, the very adaptation that made us human, and the adaptation that constitutes the latest 'major transition of life' (Maynard Smith & Szathmáry 1995). Maynard Smith & Szathmáry's other major transitions happened long ago; why was this one so tardy? Many species would seem to have rudimentary communication systems that could fall under the influence of the coevolutionary mechanism outlined above. Under what scenario might humans, but not other species, come under selection for an advanced inter-individual communication system?

The functional hallmark of human language is to combine a large vocabulary of meaningful words into utterances using syntactic rules that convey much information about how the string of words in an utterance modify each other's meanings. How might evolutionary processes favor the evolution of a high information volume communication system in humans? One possible answer is that humans were the first to acquire some sort of cognitive breakthrough. For example, Sperber (2000) is concerned with whether 'metarepresentation' (the ability to represent the thoughts of others) is a precondition for the evolution of language and Donald (2001) with the issue of whether human cultural capacities had to be fairly advanced before language could arise. In contrast, we suggest that syntax and semantics are the easy part of the evolution of language. The hard part is to figure out how humans could make use of language. Less provocatively, given innate cognitive adaptations of some kind were necessary for language, to use language, hearers must trust speakers. If they did not, they would not listen, and language would collapse. Grice (1975: 45) noted that to function conversation had to conform to a cooperative principle. But evolutionary biology teaches us that trust between individuals evolves with difficulty. Hence, the difficulty from the evolutionary perspective is how trust can evolve. If trust evolves, the cognitive bases of complex communication might evolve without serious impediment. The problem of cooperation is central to a number of accounts of the evolution human language (e.g., Hurford 2007, Tomasello 2008).

Second, how can we account for the large role of culture in language? Noam Chomsky has recently advocated a minimalist approach to the innate components of language (Chomsky 1995, 2005, Hauser *et al.* 2002). Minimalists themselves debate

about exactly what is innate, what is perhaps innate but not specific to language, and what can be off-loaded to the cultural system (e.g., Arsenijević & Hinzen, this volume). Generative grammarians who pursued the idea that surface diversity of grammars could be rigorously tied to a parsimonious set of principles and parameters seem not to have succeeded (Newmeyer 2005). Agreeing with this assessment, sociolinguists and historical linguists propose mechanisms, based on observations about how language actually changes generation by generation, that seem to account for how a vast diversity of languages can evolve culturally (Guy 1996, Deutscher 2005). Whatever is innate is very deep and hardly constrains the design space of syntax and lexicon (Progovac, this volume).

We can imagine that language learning and subsequent communication would be more efficient if more of it was innate. This is seemingly the logic of the much mooted massive modularity hypothesis for cognitive architecture generally (Tooby & Cosmides 1989, Sterelny 2003). For example, we could imagine a communication system that is largely specified innately (Sampson 2002). Such a system might superficially resemble a family of languages like the Romance Languages. Most of the grammar and much of the vocabulary would be quite similar across all the world's languages, yet a 'Castilian' child adopted by 'Catalans' might have considerable difficulty learning 'Catalan' because of having innate adaptations to speak 'Castilian'. Or perhaps everyone would just innately speak 'Catalan', but with minor culturally determined 'Valencian' and 'Barcelonan' dialects. This scenario is clearly counterfactual. Languages seem to be much more culturally diverse than they need to be for communicative efficiency. Students of the structural variation in human languages keep discovering structural principles as the number of well-studied languages grows to the point of questioning whether there are any language universals at all (Evans & Levinson 2009). Indeed, their balkanization into thousands of mutually unintelligible languages and tens of thousands mutually difficult to understand dialects is odd if selection fell only on communicative efficiency. Wouldn't communicative efficiency be maximized if we innately spoke a common regular language like Esperanto? We need to ask if language *variation* is somehow adaptive.

## 2.      Why Possibly Language is Restricted to Humans

Language is not the only rich communication system in nature. The cells in our body use nerves and hormones to coordinate our exceedingly complex development, physiology and behavior. For example, the development of complex organisms proceeds by intercellular signaling such that each cell eventually matures in the type appropriate for the tissue it finds itself in. The famous FOXP2 gene codes for a regulatory protein active in the development of all vertebrates. The common human variant differs at two amino acids from that in chimpanzees and a particular rare variant causes a specific language impairment in one well-studied family. However,

FOXP2 is expressed in many tissues during development in humans. Preliminary broad scale scans suggest that it may have hundreds of targets in a variety of tissues (Fisher & Scharff 2009). If so, one might imagine that genes exhibit the property of compositionality. FOXP2 is like a 'word' with a particular meaning that can participate in many developmental 'sentences'. We are not aware of any careful analysis intercellular communication systems for formal analogies to language, but as we learn more about how these systems work, they might be grist for comparative analysis. Interestingly, biologists have described many mechanisms by which rich internal communication systems are 'policed' to ensure a community of interest, safeguarding the basis of communication (Frank 2003). Such policing is necessary because pressure from deceptive entities, like cancer cells, to reproduce at the expense of the collective always exists.

The mathematical theory of animal communication was pioneered by John Maynard Smith (1976, 1994) and has recently been applied to human language by Lachman *et al.* (2001) and Lachmann & Bergstrom (2001, 2004). The basic problem is that individuals do not normally have enough commonality of interests to make honest communication possible. The relationship between individuals is normally competitive, and any communication thus tends to be self-interested. However, this self-interestedness means that recipients of communications cannot trust the talk of communicators, so free communication does not evolve. Thus, animal communication is normally extremely impoverished. Animal communications are costly signals, cases in which deceptive signals can be punished cheaply, or cases in which the 'communication' is an unmodifiable index of some attribute, as size is an index of strength. Peacocks' tails and the red deer stag's adaptations for male-male combat are two stock examples of costly signals. These features advertise the size and health status of males, the quality of their genes, and their fighting prowess to others in the vicinity in order to attract females and intimidate competing males. These signals are trustworthy indicators of male quality because only large, healthy males can carry a large display of gaudy feathers or develop the muscles and antlers to be competitive in fights between males. A male that produced deceptively large but weak antlers would risk injury if he had his bluff called by males with well-constructed antlers. A population might start with cheap, accurate signals of genetic quality or fighting prowess, but such communication will be undermined by mutant males that use resources to fake the signal, resulting in an arms race that ends up with an expensive unfakable signal. At an evolutionary equilibrium, the signal will be so costly that inferior males simply have insufficient resources to generate a fake signal of quality. Such a system has such a high overhead that it can produce only a handful of critical messages. Maynard Smith & Harper (2003) discuss some situations in which signals do not have to be costly subject to punishment. The most important one relevant to language is that in some situations individuals have no conflicts of interest. For example a foreign motorist might ask us to explain a traffic rule. We would have an interest in providing the correct answer, at least if the foreigner is driving on the same road, because having the correct answer will make her less likely to hit us.

Lachmann & Bergstrom show that a combinatorial communication system that relies on rules to interpret a sentence composed of words is especially vulnerable to deceptive messages and is therefore requires considerable cooperation to evolve. The value of a combinatorial communication system is that it permits limitless flexibility in the construction of messages using a finite number of words, but deceivers can use this flexibility to compose artful manipulative messages that recipients will interpret to the advantage of the signaler but to their own detriment. Merely policing the meanings of words and grammatical rules will not prevent lying. Thus, the theory tells us that rich combinatorial language-like systems evolve under a special handicap. This perhaps helps explain why language-like rich communication systems are of very restricted occurrence.

Human language is part of the extraordinary cooperation of humans with distantly related or unrelated people. The philosopher of language Paul Grice was the first person to note the centrality of cooperation in language, and many evolutionary linguists have concurred (Grice 1975; see also Sperber & Wilson 1995, Dessalles 1998, Knight 1998, Csibra & Gergely 2006, and Hurford 2007). Cooperation in language exists at two levels (Hurford 2007). First speakers must agree on common definitions of words and use the same grammatical rules if any communication is to take place at all. Second, language usage would collapse (or never arise in the first place) if speakers did not generally provide useful information to hearers, as the mathematical theory suggests. We take it that coordination on the meaning of words and grammatical rules is an easier evolutionary problem than the problem of maintaining high levels of truthfulness in the face of self-serving temptations. Because humans cooperate in large groups of distantly related and unrelated individuals, we both have a use for language — to organize our social life — and enough commonality of interest to support a rich, cheap, and trustworthy communication system (Nettle 2006, Tomasello 2008). We also police the system — telling a lie is a sin (Scott-Phillips 2008).

A number of evolutionary hypotheses are on the table to explain this cooperation (Richerson *et al.* 2003, Krebs 2008). One of these is the Tribal Social Instincts hypothesis (Richerson & Boyd 1999), which is a modernization of a hypothesis first proposed by Darwin (1874) in the *Descent of Man*. It proposes that human cooperation arose because human *cultural* variation is especially susceptible to selection at the level of the group. The ancestral groups that were relevant during the evolution of our social psychology were tribes, ethnolinguistic units ranging in size from a few hundred to a few thousand people. Normally, selection is strongest at the level of individuals, leading to competition between individuals for reproductive success. Groups of individuals can fall under strong enough selection to lead to within group cooperation, but the conditions are fairly stringent except when groups are small or closely related. Kin assortment leads to reduced genetic variation within groups and exaggerated variation between them, and group selection requires persistent genetic variation at the group level to work. The reasons that culture is unusually susceptible to group selection on a much larger scale than genetic kin groups are outlined in Richerson & Boyd (2005: 203–213) and more formally by

Henrich (2004). The linguist reader should be warned that the evolution of cooperation literature is bedeviled by terminological claims masquerading as scientific ones. Recent examples include West *et al.* (2007) and Gardner & Grafen (2009).

The basic idea is that culture much more easily than genes preserves variation between large groups of unrelated or distantly related individuals. Culture evolves more rapidly than genes because culture is modified by learned innovations not just random mutation and because we can actively choose superior variants either by direct experimentation with variants or by rules of thumb like imitating the successful or prestigious. Preserving variation between groups requires that variation between groups has to evolve fast enough to mitigate the blending effects of migration. In genetic evolution, selection and drift are generally not able to maintain much variation between neighboring groups whereas cultural processes can. Bell *et al.* (2009) used the World Values Survey to estimate cultural variation between neighboring nations and compared these estimates to published estimates of genetic variation for the same nations. Mean cultural differences are more than ten times greater than genetic differences, and the cultural differences are probably underestimated. Many important differences between groups are institutions maintained by social sanctions (Herrmann *et al.* 2008). If punishment keeps deviants rare, the differences between neighboring societies can be virtually categorical. We will shortly consider a linguistic mechanism that limits mixing between groups, helping to preserve between group variation.

Innate elements of our social psychology — Darwin argued that sympathy and patriotism were among our 'social instincts' — would have evolved by gene-culture coevolution following the same logic outlined above for the language faculty. Language would have been an important element of the evolution of cooperation. Language is obviously a general purpose communication system that today we use to talk about practically everything. For the purpose of formulating hypotheses about the intertwined evolution of language and cooperation, we need to imagine what the main function(s) of protolanguage were near the beginning of the co-evolutionary sequence that eventually gave rise to our advanced capacity for language. To get an impression of what mimetic pre-linguistic but otherwise highly imitative hominins might have been able do with without language, Donald (1991) reviewed the literature on 19th century people deaf from birth (so-called then 'deaf-mutes'). Deaf-from-birth people led surprisingly normal lives even when they acquired no language skills beyond simple 'home signs' they invented to communicate with family members. Most of them learned skills necessary to support themselves, typically a manual trade or farming, and some became prosperous. Many married and otherwise had relatively normal social lives. This evidence suggests, Donald argues, that humans might well have had a fairly advanced cultural system *before* the capacity for language evolved. Thus, cultural group selection might have selected for a measure of extra-familial altruism before the evolution of language even began.

However, we think it unlikely that social systems of the complexity that living hunter-gatherers operate could have functioned without language (see Dunbar 1996

for a slightly different version of this argument). The social life of even the simplest known hunter-gatherer societies, such as the San peoples of the Kalahari and the Shoshone of the North American Great Basin, was regulated by institutions, commonly known rules governing social behavior (Maryanski & Turner 1992). For example, the famous !Kung (Ju/'hoansi) San peoples have a complex tripartite kinship system that in effect makes everyone at least a fictive kin to everyone else, even if no genea-logical connection is known (Lee 1986). They also have a ceremonial gift exchange system that cements ties of friendship with people living in distant camps (Wiessner 2002). A large number of the gift relationships are with distantly related people, especially on the part of successful hunters and their wives. Even these comparatively simple institutions are encoded, transmitted, and reinforced linguistically through stories, gossip, complaining and shaming (Wiessner 2005). In many cases, hunter-gatherer social institutions were not at all simple. The marriage systems of the societies in the Australian central desert were very elaborate, for example (Yengoyan 1968). One might imagine that the first steps toward human cooperation in large groups might have taken place before language, but it is hard to see complex institutions arising and being maintained across a large group without it.

The general proposition that human intelligence evolved for its social functions is sometimes called the social intelligence or Machiavellian intelligence hypothesis (Dunbar 1996, Whiten & Byrne 1997). It is often contrasted with an ecological intelligence hypothesis. We think that the social intelligence hypothesis as usually stated does not give sufficient attention to the fact that hunter-gatherer *subsistence* was an intensely *social* enterprise (Steward 1955). In humans, social intelligence is fundamental to our ecological adaptation. The human hunting and gathering adaptation depended substantially upon cooperative big game hunting, especially during the late Pleistocene (Stiner 2002). Hunting probably provided the nutrients necessary to grow and sustain our very expensive brain (Aiello & Wheeler 1995, Kaplan *et al.* 2000). Many of the skills involved in hunting could perhaps have been learned by alinguistic imitation — all those skills for which a picture is worth a thousand words. However, at least one skill, tracking, is practiced in living hunter-gatherers as a sophisticated collective hypothesis testing enterprise in which verbal discussion is essential (see also Liebenberg 1990, Guthrie 2005). Trackers need to have natural-historical knowledge that often outruns that of the modern naturalists who have interacted with them (Blurton-Jones & Konner 1976) and would no doubt be hard to transmit without language.

The evolution of human intelligence might well have been driven directly by recent climate deterioration. Some theoretical and empirical evidence supports the idea that the evolution of complex, cumulatively improvable culture, including languages, was driven directly by increasingly variable environments over the last few million years (Calvin 2002, Richerson *et al.* 2005). Simple evolutionary models of culture suggest that the main adaptive advantage of a costly system of social learning is adaptation to environments with rapid, high amplitude variation (Boyd & Richerson 1985). If environmental variation is modest or if change is slow, genetic

evolution tracks changes well enough. If it is very rapid, then adaptation is via mechanisms of phenotypic flexibility like individual learning and important parts of the immune system. Combining individual phenotypic flexibility with the inheritance of acquired variation by social learning is adaptive in environments with high variability but intermediate rates of change. Culture can economize on learning by creating transmitted adaptations to environments that persist for a generation to some tens or even hundreds of generations. Work on ice and ocean cores has shown that the last ice age had the sort of high amplitude millennial and sub-millennial scale variation climate variation that the models predict would favor culture (Alley 2000). A recent ocean core suggests that the amount of millennial and sub-millennial scale variation has increased over the last four ice ages (Martrat *et al.* 2007). If the symbolic artifacts recovered by paleoanthropologists — scarcer and simple before 50,000 year ago, more numerous and more sophisticated after that date — are indicative of the emergence or improvement of languages, they may have evolved in response to the demands of the hypervariable climate of the last ice age. Coincidently or not the hypervariable part of the last ice age started about 70,000 years ago, not long before the paleoanthropological evidence suggests that human symbolic capacities and other indications of behavior likely to be related to language appeared in the record, and shortly before modern humans spread out of Africa (Richerson *et al.* 2009). As far as is currently known, the particular pattern of climate variability of the Plio-Pleistocene ice ages is unique and is thus a candidate to help explain why the human 'major transition' occurred so recently.


## 3.     Why is Language Weakly Constrained by Genes?

The number and diversity of human languages is stunning, akin to the species diversity of rain forest trees and coral reef fishes (Nettle 1999). The total number of living languages in the world is currently about 7,000 (http://www.ethnologue.com), about 1,000 of which are spoken on New Guinea. Unlike species that typically differ in functionally important ways, languages are largely functionally equivalent.

Several authors have suggested that the function of language and similar symbolic differences is to mark the boundaries of social groups (Labov 1973, Nettle & Dunbar 1997, Livingstone & Fyfe 1999, Nettle 1999, and Livingstone 2002). They serve to identify those who are 'like us' and those who belong to other groups. Language differences function to *limit* communication and hence the spread of ideas from one community to another. People from other groups are liable to tell you self-serving lies. They are also likely to tell you things that are useless in the physical and social ecology in which you live. Human cultures are ecologically a lot like species. They typically have different social institutions and often different ways of making a living (Barth 1969). Imitating people from a different culture may often be a bad thing to do (Gil-White 2005). Language diversity perhaps limits communication adaptively.

Mathematical models support this line of reasoning. Boyd & Richerson (1987) studied a cultural evolutionary model in which two populations lived in different environments. A different value of a quantitative adaptive character was optimal in each environment and there was a generic neutral marker character, also quantitative. (Quantitative characters are behaviors that can be measured on a continuous scale, such as the location of a vowel in formant space). Individuals migrated between the two populations, tending to homogenize them at some intermediate value of the adaptive character not well adapted to either environment. Counteracting mixing, juveniles were assumed to have a tendency to adopt the value of the adaptive character from people whose value of the neutral marker, say their dialect, resembled theirs. They also prefer to imitate people who are successful. In the model, the marker characters in the two populations diverged, generating a correlation between the marker and the adaptive trait. At equilibrium, the mean value of the adaptive trait was at the optimum in both environments. The preference for imitating people like you with regard to a neutral marker trait in the presence of a strong correlation between adaptive and marker characters set up an adaptive barrier limiting the flow of wrong ideas from the other environment even in the face of rather strong physical migration.

McElreath *et al.* (2003) studied a conceptually similar model but this time using discrete characters. They also modeled the environment as a social game of coordination rather than a different physical environment. Games of coordination are ones in which high payoffs depend upon matching the behavior of others. Languages are a massive game of coordination. Unless our grammar and lexicon at least roughly match those with whom we wish to communicate, we will fail. Once again migration tends to homogenize the populations but a correlation arises between the neutral marker characters and the move in the game of coordination. Eventually people in the two populations usually play with partners that correctly match their coordination move. Because language can evolve differences so rapidly, it can evolve to calibrate our discrimination against outsiders quite sensitively. I may perfectly well understand someone whose dialect differs only modestly from mine, but I may still distrust them. On the other hand, someone speaking a strange dialect is offering me something that I can judge independently to be highly useful, I might adopt it despite a general suspicion of such folks. Sociolinguists tell us that we make many social decisions based on subtle linguistic differences. Those who study the diffusion of innovations tell us that we can overcome linguistic barriers to the diffusion of ideas and techniques if we are sufficiently motivated. The models are very crude by comparison, but they lend credence to the adaptive importance of linguistic diversity but also to the flexibility of linguistic boundaries. Efferson *et al.* (2008) verified that symbolic markers could evolve in laboratory microsocieties.

Interestingly, in McElreath *et al.*'s (2003) model, symbolic differentiation cannot arise in the first instance to mark cooperators. (See Nettle & Dunbar (1997) for a model with different results.) The reason is that carrying a mark that identifies an individual as a cooperator favors the evolution of selfish non-cooperators who carry

the mark and victimize true cooperators. Games of coordination and ordinary adaptive differences do not contain a dilemma of cooperation and are not vulnerable to this problem. What symbolic differences, first among them linguistic differences, will do is inhibit communication between groups, tending to preserve variation between groups. Selection between such groups can then lead to ingroup cooperation and distrust of outgroups (Richerson & Boyd 2005).

These model results are in rough accord with the literature on ethnicity (LeVine & Campbell 1972). People speaking different languages may not normally communicate much, but, unlike biological species, they are not completely informationally isolated from one another. Ethnographic data from small-scale societies suggest that linguistic and other symbolic boundaries often mark the limits of trust and cooperation. The role of ethnicity in conflict in modern societies is well known (e.g., Horowitz 2001).

The role of language as a symbolic marker of group boundaries has been well studied by sociolinguists (Labov 1980, 2001). Labov and other sociolinguists discovered that subtle phonological changes grow up rapidly wherever social fault lines arise. For example, on Martha's Vineyard, a rural island off the coast of Massachusetts, the influx of summer tourists caused Vineyarders' speech to diverge detectably from the standard New England dialect. People seem to condition their behavior on the dialects that others speak. It is not hard to project the sociolinguistic mechanisms that cause small-scale generation to generation changes in languages to the multigeneration scale on which new dialects, new languages, and eventually new language families evolve (Nettle 1999, Deutscher 2005). Among the fields of language change that strike us as well studied include grammaticalization, the process by which conventional lexical constructions come to have syntactic functions (Hopper & Traugott 2003) and the way in which languages in contact influence each other (Thomason 2001).

## 4.    Conclusion

The theory of gene-culture coevolution has provided a useful framework for analyzing the evolution of language. Indeed, we are not aware of any contemporary students of language evolution who do not subscribe to some form of coevolutionary argument. Evolutionary linguists do differ about the division of labor between genes and culture that they propose. Some authors such as Tomasello (2008) and Kirby *et al.* (2009) posit a large role for culture in adapting language to pre-linguistic cognitive capacities or ones shared with other aspects of culture. Others, for example Pinker (2003), argue for much language specific innate cognitive adaptation. Evolutionists could contribute to this debate if it were clear what the design constraints on cognitive architecture are. The idea that specialized cognitive architecture is highly efficient is appealing, but the relatively general heuristics like joint attention that make cultural transmission efficient are impressive as well. The detailed answer

to the division of labor question is, we think, largely a matter for neurobiologists to unravel. Evolutionary considerations can provide a theory for the level of trust and cooperation necessary to make cheap, accurate, and abundant communication between humans possible and an explanation for why at least some cultural variation remains in language. In conjunction with data from paleoanthropology, paleoclimatology, and paleoecology we can produce hypotheses about when language probably evolved and why it evolved only in the human lineage. Without doubt, all of these are very difficult questions. We do not flatter ourselves that the specific proposals here will survive challenges of new data, for example data on the history of the genes that underpin language. We are more confident that the form of the coevolutionary analysis is correct.

## References

Aiello, Leslie C. & Peter Wheeler. 1995. The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Current Anthropology* 36, 199–221.

Alley, Richard B. 2000. *The Two-Mile Time Machine: Ice Cores, Abrupt Climate Change, and Our Future*. Princeton, NJ: Princeton University Press.

Arsenijević, Boban & Wolfram Hinzen. This volume. Recursion as a human universal and as a primitive.

Baptista, Luis F. & Pepper W. Trail, 1992. The role of song in the evolution of passerine birds. *Systematic Biology* 41, 242–247.

Barth, Fredrik. 1969. *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. Boston, MA: Little Brown & Company.

Bell, Adrian V., Peter J. Richerson & Richard McElreath. 2009. Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences USA* 106, 17671–17674.

Bloom, Paul. 2000. *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

Blurton-Jones, Nicholas & Melvin Konner. 1976. !Kung knowledge of animal behavior. In Richard B. Lee & Irven DeVore (eds.), *Kalahari Hunter-Gatherers: Studies of the !Kung San and Their Neighbors*, 325–348. Cambridge: Cambridge University Press.

de Boer, Bart. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28, 441–465.

Boyd, Robert & Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Boyd, Robert & Peter J. Richerson. 1987. The evolution of ethnic markers. *Cultural Anthropology* 2, 65–79.

Boyd, Robert & Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

Briscoe, Ted. 2009. What can formal or computational models tell us about how (much) language shaped the brain? In Derek Bickerton & Erös Szathámry (eds*.), Biological Foundations and the Origin of Syntax*, 369–382. Cambridge, MA: MIT Press.

Calvin, William H. 2002. *A Brain for All Seasons: Human Evolution and Abrupt Climate Change*. Chicago, IL: University of Chicago Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA: MIT Press.

Chomsky, Noam. 2005. Three factors in Language Design. *Linguistic Inquiry* 36, 1–22.

Csibra, Gergely, & György Gergely. 2006. Social learning and social cognition: The case for pedagogy. In Yuko Munakata & Mark H. Johnson (eds.), *Processes of Change in the Brain and Cognitive Development. Attention and Performance, XXI*, 249–274. Oxford: Oxford University Press.

Darwin, Charles M.A. 1874. *The Descent of Man, and Selection in Relation to Sex*. New edn., revised and augmented. New York: Hurst & Company.

Deacon, Terrence W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton.

Dediu, Dan & D. Robert Ladd. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, *ASPM* and *Microcephalin. Proceedings of the National Academy Sciences USA* 104, 10944–10949.

Dehaene, Stanislas. 2009. *Reading in the Brain: The Science and Evolution of a Human Invention*. New York: Viking.

Dessalles, Jean-Louis. 1998. Altruism, status, and the origin of relevance. In James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, 130–149. Cambridge: Cambridge University Press.

Deutscher, Guy. 2005. *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention*. New York: Henry Holt.

Donald, Merlin. 1991. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.

Donald, Merlin. 2001. *A Mind So Rare: The Evolution of Human Consciousness*. New York: W.W. Norton.

Dunbar, Robin I. M. 1996. *Grooming, Gossip and the Evolution of Language*. Cambridge MA: Harvard University Press.

Efferson, Charles, Rafael Lalive & Ernst Fehr. 2008. The coevolution of cultural groups and ingroup favoritism. *Science* 321, 1844–1849.

Evans, Nicholas & Stephen C. Levinson, 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavior and Brain Sciences* 32, 429–492.

Feldman, Marc W. & Kevin N. Laland. 1996. Gene-culture coevolutionary theory. *Trends in Ecology and Evolution* 11, 453–457.

Fisher, Simon E. & Constance Scharff. 2009. FOXP2 as a molecular window into speech and language. *Trends in Genetics* 25, 166–177.

Frank, Steven A. 2003. Perspective: Repression of competition and the evolution of cooperation. *Evolution* 57, 693–705.

Gardner, Andy & Alan Grafen. 2009. Capturing the superorganism: A formal theory of group adaptation. *Journal of Evolutionary Biology* 22, 659–671.

Gil-White, Francisco. 2005. Is ethnocentrism adaptive? An ethnographic analysis. Unpublished manuscript, University of Pennsylvania, Philadelphia. [http://www.hirhome.com/academic/Ethnocentrism.pdf.]

Gray, Russell D., Alexei J. Drummond & Simon J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323, 479–483.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & James Morgan (eds.), *Syntax and Semantics*, vol. 3: *Speech Acts*, 41-58. London: Academic Press.

Guthrie, R. Dale. 2005. *The Nature of Paleolithic Art*. Chicago, IL: Chicago University Press.

Guy, Gregory R. 1996. *Towards a Social Science of Language: Papers in Honor of William Labov*. Amsterdam: John Benjamins.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Henrich, Joe & Richard McElreath. 2008. Dual inheritance theory: The evolution of human cultural capacities and cultural evolution. In Robin Dunbar & Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology*, 571–585. Oxford: Oxford University Press.

Henrich, Joseph. 2004. Cultural group selection, coevolutionary processes and large-scale cooperation. A target article with commentary. *Journal of Economic Behavior and Organization* 53, 3–143.

Herrmann, Benedikt, Christian Thöni & Simon Gächter. 2008. Antisocial punishment across societies. *Science* 319, 1362–1367.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Horowitz, Donald L. 2001. *The Deadly Ethnic Riot*. Berkeley, CA: University of California Press.

Hurford, James R. 2007. *The Origins of Meaning*. Oxford: Oxford University Press.

Kaplan, Hillard, Kim Hill, Jane Lancaster & A. Magdalena Hurtado. 2000. A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology* 9, 156–185.

Kirby, Simon, Morton H. Christiansen & Nick Chater. 2009. Syntax as an adaptation to the learner. In Derek Bickerton & Eros Szathmary (eds.), *Biological Foundations and the Evolution of Syntax*, 325–343. Cambridge, MA: MIT.

Knight, Chris. 1998. Ritual/speech coevolution: A solution to the problem of deception. In James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, 68–91.

Cambridge: Cambridge University Press.

Krebs, Dennis L. 2008. Morality: An evolutionary account. *Perspectives on Psychological Science* 3, 149–172.

Labov, William. 1973. *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

Labov, William. 1980. *Locating Language in Time and Space*. New York: Academic Press.

Labov, William. 1994. *Internal Factors — Principles of Linguistic Change*, vol. I. Malden, MA: Blackwell.

Labov, William. 2001. *Social Factors — Principles of Linguistic Change*, vol. II. Malden, MA: Blackwell.

Lachmann, Michael & Carl T. Bergstrom. 2004. The disadvantage of combinatorial communication. *Proceedings of the Royal Society of London B* 271, 2337–2343.

Lachmann, Michael, Szabolcs Szamado & Carl T. Bergstrom. 2001. Cost and conflict in animal signals and human language. *Proceedings National Academy of Sciences USA* 98, 13189–13194.

Lee, Richard. 1986. !Kung kin terms, the name relationship and the process of discovery. In Megan Biesele, Robert Gordon & Richard Lee (eds.), *The Past and Future of !Kung Ethnography: Critical Reflections and Symbolic Perspectives — Essays in Honour of Lorna Marshall*, 77–102. Hamburg: Helmut Buske Verlag.

LeVine, Robert & Donald T. Campbell. 1972. *Ethnocentrism: Theories of Conflict, Ethnic Attitudes, and Group Behavior*. New York: Wiley.

Liebenberg, L. 1990. *The Art of Tracking: The Origin of Science*. Cape Town: David Phillip.

Livingstone, Daniel. 2002. The evolution of dialect diversity. In Angelo Cangelosi & Domenico Parisi (eds.), *Simulating the Evolution of Language*, 99–117. London: Springer.

Livingstone, Daniel & Colin Fyfe. 1999. Modeling the evolution of linguistic diversity. In Dario Floreano, Jean Daniel Nicoud & Francesco Monada (eds.), *Proceedings of the 5th European Conference in Artificial Life: ECAL '99*, 704–708. Lausanne: Springer.

Martrat, Belen, Joan O. Grimalt, Nicholas J. Shackleton, Lucia de Abreu, Manuel A. Hutterli & Thomas F. Stocker. 2007. Four climate cycles of recurring deep and surface water destabilizations on the Iberian margin. *Science* 317, 502–507.

Maryanski, Alexandra & Jonathan H. Turner. 1992. *The Social Cage: Human Nature and the Evolution of Society*. Stanford, CA: Stanford University Press.

Maynard Smith, John. 1976. Sexual selection and the handicap principle. *Journal of Theoretical Biology* 57, 239–242.

Maynard Smith, John. 1994. Must reliable signals always be costly? *Animal Behaviour* 47, 1115–1120.

Maynard Smith, John & David Harper. 2003. *Animal Signals*. Oxford: Oxford University Press.

Maynard Smith, John & Eörs Szathmáry. 1995. *The Major Transitions in Evolution*.

Oxford: W.H. Freeman / Spectrum.

McElreath, Richard, Robert Boyd & Peter J. Richerson. 2003. Shared norms and the evolution of ethnic markers. *Current Anthropology* 44, 122–129.

Munroe, Steve & Angelo Cangelosi. 2002. Learning and the evolution of language. *Artificial Life* 8, 311–339.

Nettle, Daniel & Robin I. M. Dunbar. 1997. Social markers and the evolution of reciprocal exchange. *Current Anthropology* 38, 93–99.

Nettle, Daniel. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.

Nettle, Daniel. 2006. Language: Costs and benefits of a specialized system for social information transmission. In Jonathan C. K. Wells, Simon Strickland & Kevin Laland (eds.), *Social Information Transmission and Human Biology*, 137–152. London: Taylor & Francis.

Nettle, Daniel. 2007. Languages and genes: A new perspective on the origins of human cultural diversity. *Proceedings National Academy of Sciences USA* 104, 10755–10756.

Newmeyer, Fredrick J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Origgi, Gloria & Dan Sperber. 2000. Evolution, communication and the proper function of language. In Peter Carruthers & Andrew Chamberlin (eds.), *Evolution and the Human Mind: Modularity, Language, and Meta-Cognition*, 140–169. Cambridge: Cambridge University Press.

Pagel, Mark. 2009. Human language as a culturally transmitted replicator. *Nature Reviews Genetics* 10, 405–415.

Pepperberg, Irene M. 2000. *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Cambridge, MA: Harvard University Press.

Pinker, Steven. 2003. Language as an adaptation to the cognitive niche. In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution*, 16–37. Oxford: Oxford University Press.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.

Progovac, Ljiljana. This volume. Syntax: Its evolution and its representation in the brain.

Richerson, Peter J. & Robert Boyd. 1989. The role of evolved predispositions in cultural evolution: Or sociobiology meets Pascal's Wager. *Ethology and Sociobiology* 10, 195–219.

Richerson, Peter J. & Robert Boyd. 1999. Complex societies — The evolutionary origins of a crude superorganism. *Human Nature* 10, 253–289.

Richerson, Peter J., Robert Boyd & Joseph Henrich. 2003. Cultural evolution of human cooperation. In Peter Hammerstein (ed.), *Genetic and Cultural Evolution of Cooperation*, 357–388. Cambridge, MA: MIT Press.

Richerson, Peter J., Robert L. Bettinger & Robert Boyd. 2005. Evolution on a restless planet: Were environmental variability and environmental change major drivers of human evolution? In Franz M. Wuketits & Francisco J. Ayala (eds.),

*Handbook of Evolution: Evolution of Living Systems (Including Hominids)*, 223–242. Weinheim: Wiley-VCH.

Richerson, Peter J. & Robert Boyd. 2005. *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: University of Chicago Press.

Richerson, Peter J., Robert Boyd & Robert L. Bettinger. 2009. Cultural innovations and demographic change. *Human Biology* 81, 211–235.

Sampson, Gregory. 2002. Exploring the poverty of the stimulus. *The Linguistic Review* 19, 73–104.

Savage-Rumbaugh, E. Sue & Roger Lewin. 1994. *Kanzi: The Ape at the Brink of the Human Mind*. New York: Wiley.

Scott-Phillips, Thomas C. 2008. On the correct application of animal signalling theory to human communication. In Andrew D. M. Smith, Kenny Smith & Ramon Ferrer i Cancho (eds.), *The Evolution of Language: Proceedings of the 7Th International Conference (EVOLANG7)*, 275–282. Singapore: World Scientific.

Smith, Kenny & Simon Kirby. 2008. Cultural evolution: Implications for the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B* 363, 3591–3603.

Sperber, Dan. 2000. Metapresentations in an Evolutionary Perspective. In Dan Sperber (ed.), *Metarepresentations: A Multidisciplinary Perspective*, 117–137. Oxford: Oxford University Press.

Sperber, Dan & Deirdre Wilson. 1995. *Relevance: Communication and Cognition*, 2nd edn. Malden, MA: Blackwell.

Steels, Luc. 2009. Cognition and social dynamics play a major role in the formation of grammar. In Derek Bickerton & Eörs Szathmáry (eds.), *Biological Foundations and the Origin of Syntax*, 345–368. Cambridge, MA: MIT.

Sterelny, Kim. 2003. *Thought in a Hostile World: The Evolution of Human Cognition*. Malden, MA: Blackwell Publishing.

Steward, Julian H. 1955. *Theory of Culture Change: The Methodology of Multilinear Evolution*. Urbana, IL: University of Illinois Press.

Stiner, Mary C. 2002. Carnivory, coevolution, and the geographic spread of the genus homo. *Journal of Archaeological Research* 10, 1–63.

Thomason, Sarah Grey. 2001. *Language Contact*. Washington, DC: Georgetown University Press.

Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tooby, John & Leda Cosmides. 1989. Evolutionary psychology and the generation of culture — 1. Theoretical considerations. *Ethology and Sociobiology* 10, 29–49.

West, Stuart A., Ashleigh S. Griffin & Andy Gardner. 2007. Social semantics: How useful has group selection been? *Journal of Evolutionary Biology* 21, 374–385.

Whiten, Andrew & Richard W. Byrne. 1997. *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge: Cambridge University Press.

Wiessner, Polly. 2002. Hunting, healing, and hxaro exchange: A long-term perspective on !Kung (Juꞌhoansi) large-game hunting. *Evolution and Human*

*Behavior* 23, 407–436.

Wiessner, Polly. 2005. Norm enforcement among the Ju/'hoansi Bushmen: A case for strong reciprocity? *Human Nature* 16, 115–145.

Yengoyan, Aram A. 1968. Demographic and ecological influences on aboriginal Australian marriage systems. In Richard B. Lee & Irven DeVore (eds.), *Man the Hunter*, 185–199. Chicago, IL: Aldine.

*Peter J. Richerson*                                                  *Robert Boyd*
*University of California, Davis*                              *University of California, Los Angeles*
*Department of Environmental Science & Policy*   *Department of Anthropology*
*One Shields Avenue*                                             *375 Portola Place*
*Davis, CA 95616*                                                  *Los Angeles, CA 90095*
*USA*                                                                       *USA*
*pjricherson@ucdavis.edu*                                    *rboyd@anthro.ucla.edu*

**TABLE OF CONTENTS**

★ EDITORIAL ★