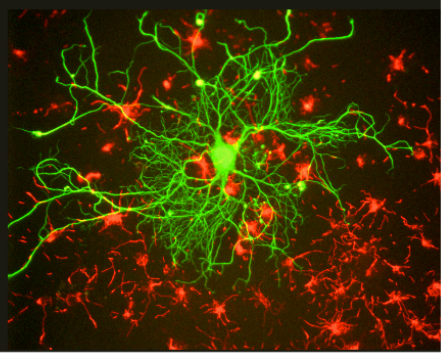
Advances in Biolinguistics

25-27 July 2013, Geneva



A workshop organized by
the International Network in Biolinguistics

Abstracts

Advances in Biolinguistics

ICL19 International Congress of Linguists

A workshop organized by
the International Network in Biolinguistics

Abstracts

Contents

Logic and Language

Author: Roland FRIEDRICH, Humboldt-Universität, Berlin

Abstract: The question of how the language faculty and reasoning are related is not yet fully understood. The strongest hypothesis states that reasoning is fully embedded into language, i.e., the core computational procedures which subserve natural language processing are also the ones that subserve reasoning, logical thinking and mathematics. However, there is increasing evidence, e.g., based on fMRI data, that language does not provide the sought after substrate but rather has itself to be understood as emerging from the interaction of more basic macroscopic neural networks. In this talk we shall review this question from a network centric perspective and present related experimental findings.

Bioinformatic properties of sign language motion:

fractal complexity of optical flow

Author: Evie MALAIA, University of Texas at Arlington
Co-Authors: Joshua D. BORNEMAN, Purdue University
Ronnie B. WILBUR, Purdue University

Abstract: A fundamental goal in sign language research is to identify the perceptual properties of the visual signal that underlie the linguistic universals in human languages. In the auditory domain, world languages and music have been described as having modulation spectra of moderate fractal complexity (1/f) (Singh & Theunissen, 2003). Perceptual research (Yu, Romero, & Lee, 2005) shows that visual (V1) neurons in the human brain are tuned to optimally respond to 1/f complexity of visual signal, suggesting a biological basis for neural sensitivity to moderate fractal complexity in visual stimuli. We hypothesized that in American Sign Language, the distinctive grammatical categories will fall within the range of moderate complexity with regard to their kinematic variability. To test this, we analyzed frequency profiles of optical flow for two types of videos: 1) natural scenes containing humans conducting everyday activities (folding laundry, assembling a Playstation), and 2) signers of American Sign Language (ASL) producing grammatically distinct signs. The frequency profile for each of the binned magnitudes of optical flow in ASL vs. human activity videos indicates more level distribution of high frequencies in sign language videos, suggesting higher fractal complexity of signing as compared to everyday human motion (Figure 1, C: lower ß value in 1/fß correspond to more fractal nature of the signal). Further research should investigate relevance of visual-kinematic properties of sign language to neurobiological mechanics of attention in the visual domain, and sign language acquisition. References Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. The Journal of the Acoustical Society of America, 114, 3394. Yu, Y., Romero, R., & Lee, T. S. (2005). Preference of sensory neural coding for 1/f signals. Physical review letters, 94(10), 108103.

Variation, speciation and the development

of sapiens' language ready brain

Author: Antonio BENÍTEZ-BURRACO, Universidad de Oviedo, Spain
Co-Authors: Lluís BARCELÓ-COBLIJN, Universitat de les Illes Balears, Spain
Cedric BOECKX, ICREA/Universitat de Barcelona, Spain

Abstract: In this talk we would like to consider the question of whether other hominins closely related to Homo sapiens had a linguistic system similar to ours. This will offer us an opportunity to reconsider the relation between genetics and the language faculty and try to explore an important issue: does having a particular set of gene(s) (e.g. FOXP2 or its interactome) imply having the Language Faculty? It is known that anatomically-modern humans (henceforth, AMH) have interbred with at least three different hominins. Does this fact imply that they also had modern language? We do not think so, and the reasons are related to natural laws and, we suggest, organic developmental patterns that can be found in nature. We claim that only AMH. had a full blown language system (though we have no problem claiming that other species could vocalize), and put forth the idea that differences accounting for this specific trait in AMH lie in developmental trajectories. This in turn leads us to call for a renewed interest in grounding theoretical claims (such as Universal Grammar) into facts about the brain. Here we rely on a specific model advanced by Boeckx (Boeckx, C. [2012] Homo combinans. Keynote talk, EVOLANG9, Kyoto), who takes the core structuring of human language syntax to be a clock-wavefront mechanism (part of a larger family of reaction-diffusion, or more accurately, Local Autoactivation Lateral Inhibition/LALI mechanisms), implemented in the brain in terms of a cortico-thalamic network (whose externalization is linked to Broca's region and the basal ganglia). The implication of the thalamus is especially significant in light of the attested sapiens-specific skull, and by hypothesis, brain globularity, achieved through development, as the position of the thalamus in a globular brain offers the possibility of more efficient connection and information exchange. We take this to be a special instance of a more general argument in favor of computational efficiency in the context of language. In sum, if AMH, Neanderthals, and Denisovans share a 99.5% of the genome, how can we argue in favor of differentiated cognitions? The answer largely lays in 1) the natural principles that regulate the brain growth and 2) the brain activity. Genetics has long provided arguments that two organisms can have the same genes and still show different ontogenic developments. We argue that – without preventing interbreeding– a differing developmental trajectory separated sapiens from Neanderthals, giving rise to distinct cognitive profiles. In other words, as Richard Lewontin would have it, even with having the genes, 'it ain't necessarily so'.

Toward historical biolinguistics

Author: Giuseppe LONGOBARDI, University of York

Abstract: Beyond its theoretical success, the development of molecular biology has brought about the possibility of extraordinary progress in the historical study of classification and distribution of different species and different human populations, introducing a new level of evidence (molecular genetic markers) now susceptible to quantitative and computational treatment. I argue that, even in the cognitive sciences, purely theoretical progress in a certain

discipline, such as linguistics, may have analogous historical impact, equally contributing to Renfrew's 'New Synthesis', and in turn may be confirmed by such results. So, exactly on the model of molecular biology and its fruitful balance between evolutionary and theoretical concerns, I propose to unify two traditionally unrelated lines of investigation: 1) the formal study of syntactic variation (parameter theory) in the biolinguistic program 2) the reconstruction of relatedness among languages (phylogenetic taxonomy) I argue that due to progress in parametric grammatical theories, and relying on the methodological parallelism with evolutionary genetics, we are now in a position to measure the syntactic distance among languages in a precise fashion and to explore its historical significance through the application of clustering algorithms borrowed from computational biology. Through such formal methods, it will be shown that the distribution of actual syntactic distances provided by an elaborate parametric system is statistically significant (that is, non-accidental and requiring historical explanation) and empirically matches well assessed historical expectations. The properties commonly attributed to parametric variation (uniform discreteness and universality), akin to those displayed by genetic polymorphisms, make it particularly suitable for addressing the unsolved issue of comparison between remote, etymologically unrelated languages, opening the way for completely new interactions with the empirical results of molecular anthropologists. Further experiments suggest that pattern-based classifications are phylogenetically less accurate than those founded on deductively structured parameters, so that the latter better represent actual historical relations. Therefore, abstract parametric theories are likely to encode a higher level of reality than surface-oriented typologies. Thus, I suggest 3) that a parametric model of the language faculty and language acquisition/transmission (more broadly a theory of generative grammar) receives original support from its historical adequacy; 4) that through these new tools one can explore the possibility of testing Darwin's (1859) prediction that, when properly identified, the genealogical trees of human populations and of their languages should eventually turn out strictly parallel.

The robustness of language development:

the role of crosspopulation and crosslinguistic studies

Author: Joao COSTA, Universidade Nova de Lisboa

Abstract: The literature on language acquisition from the last two decades reveals that most syntactic operations and configurations are acquired very early on, and that parameter values are set in the first year of life. These findings provide a good piece of evidence in favor of the innate nature of language. It is however known that not all syntactic knowledge is acquired so early, and interesting asymmetries can be observed between very similar constructions. For instance, in A-bar dependencies, object dependencies are acquired later than subject dependencies, and in the domain of referential dependencies, pronouns are harder to master than anaphors (Chien and Wexler 1990). A detailed analysis of the types of dependencies that develop slower reveals the relevance of fine-grained linguistic investigation (e.g. Friedmann, Belletti and Rizzi 2009 show that object dependencies are only problematic in intervention configurations). Furthermore, crosslinguistic variation is found: for example, pronouns do not pose problems for comprehension when they are clitic (McKee 1992, Padilla 1994). Variation in the pace of acquisition and crosslinguistic variation in acquisition provides a very interesting window into the universality of language development, but, as I will argue, do not challenge its robustness. In this talk, I aim at discussing the importance of crosslinguistic and

crosspopulation studies in order to determine what can be explained in language development in terms of fundamental language knowledge, language-specific characteristics or general processing capacities non specific to language. Crosspopulation studies: I will show that certain difficulties in language development – particularly in the parsing of movement and referential dependencies – do not follow from a lack of language knowledge, but from different working memory capacities (for this, I will report on studies comparing typically developing children, SLI children, deaf children, and patients with Alzheimer's disease). The point to be made is that certain tendencies found in children are also found in populations without any language disorder. This argues for the robustness of linguistic knowledge, and the slower development can be independently motivated. Crosslinguistic studies: I will show that difficulties found in certain languages in the development of specific constructions are not universal, and follow from language-specific properties, and that the crosslinguistic asymmetries can be explained once the role of interfaces is taken into consideration.

Should we expect a variable Faculty of Language?

Author: Antonio BENÍTEZ-BURRACO, Universidad de Oviedo, Spain
Co-Author: Cedric BOECKX, ICREA/Universitat de Barcelona, Spain

Abstract: Current psycholinguistic, neurobiological and genetic research has greatly increased the degree of variation regarding language and linguistic phenomena. In particular, it seems to cast doubt on the purportedly homogeneous nature of the language faculty. For instance, psycholinguistic measures are variable across the normal population, suggesting a variable competence/performance within it. At the brain level the boundaries of the 'language areas' are rather changeable among the diverse individuals, but also across development. Moreover, many genes contribute to regulate the development and the functioning of this neural substrate, but they are (highly) polymorphic, with some variants giving rise to pathological conditions, but with others being present as well within the unaffected population. This seems to challenge, in particular, the longstanding assumption that the linguistic genotype is going to be uniform across the species in the absence of a fairly severe and specific pathology. In this presentation we discuss whether (and to which extent) this genetic diversity can actually be reconciled with the widespread view of the faculty of language as a "one component of the human mind", in essence, as an idiosyncratic cognitive capacity/entity/ability, qualitatively equal in all human beings. In addressing this question, we appeal to (and explore the implications of) some fresh hypotheses posited by evolutionary developmental biology (Evo-Devo). In particular, we argue that developmental dynamics (and hence, an assorted set of regulatory factors) strongly canalizes variation, to the extent that the same phenotype can robustly emerge at the term of growth from diverse genotypes. We argue that language disorders can be construed as conditions for which canalization has been unable to achieve particular degrees of development. Breakdowns do not occur randomly, clearly because adaptability is always constrained, but plausibly also because certain cognitive processes are more vulnerable than others to damage or to developmental disturbances. Eventually, even though any of its biological components can be regarded as specifically linguistic, the language faculty itself can actually be characterised as an idiosyncratic cognitive faculty, almost certainly because of that pervasive tendency of their components to interface whenever growth takes place in the presence of a suitable amount of linguistic stimuli.

Natural Laws and Biolinguistics - the case of symmetry

Author: Lyle JENKINS, Boston Biolinguistics Institute

Abstract: "Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection." (Weyl, Symmetry, 1983:5) Throughout the natural sciences, principles such as least action, symmetry, stochastic processes, etc. have had a profound effect on our understanding of the world around us. An area of great interest in biolinguistic research is to what degree properties of human language might be determined by principles that originate outside the language faculty itself, similar to those mentioned above, and others. For example, it has been proposed that the computational system of language might be shaped in part by principles of efficient computation or of least effort that may share properties found in other cognitive and physical systems. We will consider the specific case of principles of symmetry which have had a far-reaching effect in unifying areas of mathematics, physics, chemistry, and more recently, some areas of biology; e.g., pattern formation. We ask in what ways the concepts of symmetry and symmetry breaking might provide some insight into linguistic phenomena. In conclusion, we will consider how these principles could serve to unify parts of linguistics with other biological systems and, more widely, the natural sciences (the "unification problem").

Evo-devo language universals

Author: Anna-Maria DI SCIULLO, Université du Québec à Montréal
Co-Authors: Marco NICOLIS, Université du Québec à Montréal
Stanca SOMESFALEAN, Université du Québec à Montréal

Abstract: Evolutionary developmental universals emerge in the historical development of languages. One such principle is the gradual elimination of the oscillation of a dependent with respect to its head (Di Sciullo 2011, 2012, 2013, Di Sciullo and Nicolis 2012, Di Sciullo and Somesfalean 2012). We provide further arguments for such symmetry breaking universals and we raise the question whether the growth of language in the individual (ontogeny) recapitulates the historical development of languages (phylogeny). We focus on the relative position of P and its pronominal complement. While cases of fluctuating asymmetry (oscillation) in the position of P heads with respect to their complement are attested in Old English (with most Ps) and in Old Italian (commitative P only), no syntactic oscillation is found in Modern English and Italian. Why would there be a bias favoring prepositions instead of postpositions in the languages under consideration? Why is there no trace of syntactic oscillation in English and Italian? And why are there no errors in the relevant contexts in CHILDES database? The oscillation between pre- and post- nominal positions wrt a functional head may follow from the availability of both a valued and an unvalued [D] feature for the functional heads F and P dominating D. Only the [uD] feature of F is checked in the case of pre-positions, whereas the [uD] feature of P would also be checked in the case of post-positions. According to the Unique Checking Constraint (UCC) (Wexler 1998), the genetic system development has the property that more than one checking is dispreferred at young ages. In the clausal domain, AGReement and Tense cannot be both checked, via the UCC, and thus one is eliminated giving rise to the Optional Infinitive stage. Similarly, in the nominal domain, postpositional structures would be more complex, and thus would be dispreferred by the genetic system. Development is at the core of both biology and language viewed as a biological

phenomenon. In this perspective, the principles of language growth are based on genetic endowment, in addition to the existence of learning processes (Chomsky 1968, 2005, 2011, Lenneberg 1967). The principles of language growth are not learned on the basis of quantitative data a child is exposed to in a given span of time. If these principles were genetically determined, it would be unlikely that language ontogeny would recapitulate language phylogeny. Based on our targeted analysis, we find no evidence supporting the recapitulation theory, whereas we find reasons to relate principles of language growth to Evolutionary developmental universals.

Context representation retrieval and update and the nature of the syntax-semantics interface

Author: Boban ARSENIJEVIC, CCSIC Madrid

Abstract: A view of language is presented in which syntax is an algorithm for the retrieval and update of a relevant context representation. Our cognitive apparatus builds, maintains and retrieves rich cross-modular context representations that can be seen as webs of referents (connected by their mutual relations). A referent is no more than a set of relations with a number of other such sets (i.e. with a number of other referents). A linguistic expression is semantically interpretable if it specifies a set of relations (each specified for its second argument), and the relevant context representation contains a referent represented as a superset of the set specified by the expression. I examine the role of syntax in the architecture of language taken this perspective.

Mental representation of natural numbers and acquisition of numerals

Author: Akira WATANABE, University of Tokyo

Abstract: In this talk, I suggest that the Merge-based conception of natural numbers advocated by Chomsky (2008) enables us to explain why full mastering of the meaning of numerals by children is a prolonged process, completed as late as at age 3½ or 4 (Carey 2009, Wynn 1992). The difficulty for children can be reduced to the unusual complexity of linguistic representations of numerals, once Chomsky's original proposal is slightly modified. [Linguistic Representations]: Chomsky (2008) suggests that repeated application of Merge to a single lexical item LI gives rise to a sequence of natural numbers as in (1). (1) 1 = one, 2 = {one}, 3 = {one, {one}}, ... Since any lexical item will do, (1) can be rendered more generally as: (1') 1 = LI, 2 = {LI}, 3 = {LI, {LI}}, ... My proposal is to shift the sequence by one unit and adopt the following: (2) 1 = {LI}, 2 = {LI, {LI}}, 3 = {LI, {LI, {LI}}}, ... In other words, one application of Merge is equated with 1, two applications with 2, and so on. The point of this modification is to match the linearized string of n LI's with the corresponding numeral list. In the case of 3, we have: (3) LI + LI + LI = One, two, three. The idea is to subject the set-theoretic object created by Merge to PF linearization, just as in the case of ordinary syntactic objects. In fact, when we count, we can insert a sequence of numerals in the DP-internal slot for a numeral as in (4). (4) We have one, two three, four, five syntacticians here. Thus, the lexical entries for numerals look like the following: (5) a. original conception (meaning): 1 = {LI}, 2 = {LI, {LI}}, 3 = {LI, {LI, {LI}}} b. linearization of the original structure (PF1): 1 = LI, 2 = LI + LI 3 = LI + LI + LI c. actual

counting sequence (PF2): 1 = one, 2 = one + two 3 = one + two + three d. actual PF (PF3): 1 = one, 2 = two, 3 = three (5b) is not used as such, but helps connect (5a) to (5c). Inclusion of (5c) as a piece of PF information derives as a lexical property one of Gelman and Gallistel's (1978) counting principles, which says that the last numeral in the sequence represents the cardinality of the set counted. [Acquisition]: Compared with ordinary lexical items, which are associated with a single phonological form, entries like (5) are unusually rich. It is quite natural to expect this complexity to slow down the acquisition process. Crucially, experimental measures used to assess children's understanding of natural number concepts (Le Corre et al. 2006) all involve expressions like "three frogs". What is difficult for children is not natural number concepts but numerals with language-particular phonological forms.

Variable elimination in language system and mathematics

Author: Koji ARIKAWA, St. Andrew's University

Abstract: Alfred Russel Wallace (1823–1913), the co-founder of the theory of evolution, was puzzled by the fact that the "gigantic development of the mathematical capacity is wholly unexplained by the theory of natural selection." Noam Chomsky suggests that the mathematical capacity is derivative from language. Gaussian elimination in the computational procedure of mathematics (CM) is the most elementary algorithm for solving simultaneous linear equations such as (1) ax + by = c and (2) dx + ey = f. I propose that CM has AGREE (MATCH + VALUE + DELETE) as CHL (contra Chomsky). To eliminate the variable x in (2), we first eliminate the coefficient d. We seek a multiplier M (matchmaker) that satisfies d – aM = 0. This 0 becomes the new coefficient for x in (2). Since 0x = 0, x is eliminated. Eliminating x, values y ("irreflexive" elimination). There is no need to eliminate every x ("partial" elimination). Back substituting y's value, we get x's value. In CHL, there are three types of uF: (i) uPHI-set in V/v* and T/C, (ii) CASE in DP, and (iii) EPP in heads. uPHI has an interpretable matching counterpart PHI in DP; it is valued by PHI ("direct" back substitution?) and deleted because of redundancy ("reflexive" elimination; look-ahead?) When uPHI is deleted, every relevant uPHI is deleted ("total" elimination). CASE does not have an interpretable matching counterpart, but is valued by a head ("direct" back substitution?) and deleted as a reflex (?) of uPHI-deletion. EPP does not have an interpretable matching counterpart and is not valued; it is deleted because of FI (look-ahead?). Suppose that CHL and CM share an elimination process. Consider the v*P phase with a two-place transitive V, leaving aside EPP (coefficient?). CHL solves the simultaneous equations: (1') auPHI + bCASE = c and (2') duPHI + eCASE = f. To eliminate uPHI in (2'), we first eliminate d. We seek M such that d – aM = 0. Suppose that M = PHI in object DP; then d – aPHI = 0. This 0 becomes the new coefficient for uPHI in (2'). Since 0uPHI = 0, uPHI is eliminated. Eliminating uPHI, values CASE, which is a linear algebraic expression of CASE as a "reflex" of uPHI-deletion. Elimination in CM fails when the elimination of d in (2) produces 0y = f', where f' ≠ 0. The two lines (each expressing (1) and (2)) become parallel, i.e., no intersection (solution). This corresponds to (2'), where d = e = 0 and both uPHI and CASE are eliminated. A candidate head is defective V (unaccusative/passive/participle) with incomplete-PHI, which does not enter into the agreement process (enables successive cyclic movement). Elimination in CM may produce infinitely many solutions if (1) = (2). The two lines pile up. This corresponds to (1') = (2'). A candidate is EPP, which is satisfied by anything and in any way (obligatorily or optionally).

The great leap forward and the emergence of complex numerals

Author: Anna-Maria DI SCIULLO, Université du Québec à Montréal

Abstract: I argue that the ability for the human mind to compute complex numerals is a consequence of the great leap from finite and continuous systems, such as the gestural system, to systems of discrete infinity, such as language, mathematics and music. I further argue that the sub-systems of the brain sub-serving grammar on the one hand, and mathematics on the other provide complementary dimensions of complex numerals. 1. The set of complex numerals is unbounded (Zwicky1963, Brainerd 1971, Radzinsky 1991, contra Merrifield 1968, Greenberg 1978, a.o.). This set is derived by Merge, the central operator of the Language Faculty, a dyadic, recursive and unbounded operation, deriving ordered pairs, according to Chomsky (2008). I focus on the recursive properties of coordinate and prepositional part-whole structures in complex numerals from different languages, where F is a functional head (conjunction, preposition, cardinal affix, Case) and Num is a lexical or a complex numeral. Simplex and complex numerals merge with a functional head at each step of the derivation, as proposed in Di Sciullo (2012). This is expected in a model where two maximal projections may never merge directly, but only indirectly via a functional projection. 2. The sub-system of the brain sub-serving mathematical reasoning provides a computation of complex numerals. I argue that while the cardinality of a complex numeral is located at the periphery of each sub-constituent, the arithmetic operators (+, x) are legible at the language-mathematics interface as features of the functional heads F asymmetrically relating the parts of complex numerals. 3. The derivation of complex numerals by the Language Faculty is subject to principles of efficient computation, such as phases and 'pronounce the minimum' Chomsky (2005, 2011). I argue further that principles of symmetry breaking are also part of these principles, ensuring that only the configurations whose components are asymmetrically related are legible by the external systems. 4. Brain imaging results indicate that processing hierarchically structured mathematical formulae and processing complex syntactic hierarchies in language activate different areas of the brain (Friederici et al 2011, Friedrich & Friederici 2009). The fact that complex numerals are composed of asymmetrical substructures headed by F heads suggests that they are processed by the part of the brain that sub-serves language. The fact that these F heads correspond to arithmetic operators, even if not pronounced, indicates that the algebraic computation of complex numerals is provided by the part of the brain that sub-serves mathematical reasoning.