UNIVERSAL PARSING AND GRAMMATICAL VARIATION

NINO GRILLO

1. Abstract

A natural assumption in linguistics is that the human language parser is governed by universal principles of economy, managing the complexity of language processing by optimizing the computations. This project deals with a large body of findings that questioned the universality of parsing principles, and particularly of Locality, a principle shown to apply to both representations and processing. The project presents an account of these exceptions based on the discovery, by the applicant, of a grammatical confound in the previous literature. Some languages allow an additional grammatical structure (Pseudo Relative) in the relevant sentences, and there is a large correspondence between the availability of this structure and the exceptions to locality. The grammatical properties of this structure clarify the epiphenomenal nature of the exceptions. The project combines tools from theoretical linguistics with behavioural (eye-tracking) and neuroimaging (ERP) methods to further our understanding of universal constraints on processing and representation and their implementation in the brain. The proposed project will further the ongoing collaboration between the applicant, Dr. Grillo and the Centro de Linguistica da Universidade Nova de Lisboa and presents an ideal blend of skills and knowledge. The researcher has substantial skills in experimental methods and knowledge in psycho/neurolinguistics, whereas the department hosts a number of leading experts in theoretical linguistics and language acquisition. As demonstrated by the applicant’s record of publications, presentations and research grants, this ongoing collaboration has been extremely fruitful for both parties. The project is expected to have a positive, lasting impact on the fields of theoretical and experimental linguistics and lead to fruitful long-term collaborations between these connected disciplines. The project will provide me with the opportunity to acquire knowledge and skills critical for my professional development and to transfer them to the CLUNL, in harmony with what already accomplished in the past years (2 MA and 1 PhD theses completed). I will strengthen my knowledge in formal linguistics and train in two new experimental methods, eye-tracking and ERP. These techniques not only are ideal to deal with the issues raised in this project, they will also enrich me as psycholinguist and neurolinguist, as I already possess the theoretical background necessary to acquire them relatively quickly.
2. B. Full description of the application

2.1. Background. Contrasts in grammaticality underpin an important hypothesis in theoretical linguistics: that syntactic representations must be built following a principle known as Locality. According to this principle, a syntactic relation initiated by a node N has to be established with the closest possible element capable of satisfying Ns needs; unnecessarily distant relationships give rise to ungrammaticality [19]. Similar results have been found in Psycho and Neurolinguistics: the human parser operates according to locality principles that govern both structure building (as evidenced by preferences in the attachment of temporal modifiers) and filler-gap processes (as evidenced by active filler effects) [10, 11].

There is one domain of psycholinguistic research, however, in which the notion of locality has been claimed to yield incorrect results: the attachment of Relative Clauses (RCs) in globally ambiguous contexts. Cuetos and Mitchell [6] found that in such contexts English speakers have a preference for Low Attachment (LA) of RCs, in line with locality (or Late Closure, as the relevant principle is known in the parsing literature) (1-a). Spanish speakers, however, preferred High Attachment (HA), apparently favouring non-local over local attachment (1-b). Successive studies have confirmed the reality of this asymmetry using a variety of offline and online experimental methods (including eye-tracking and ERP). LA preference is found not only in English, but in Romanian, Basque and Chinese; a preference for HA was reported not only for Spanish, but for Dutch, French, Italian, Russian and Greek.

(1) a. Someone shot the maid1 of the actress2 that was2 standing on the balcony
   b. Alguien dispar contra la criada1 de la actriz2 que estaba1 en el balcón

The typological picture is complicated by the fact that there is variation within as well as across languages. In certain syntactic environments [12]. In addition attachment has been shown to be sensitive to factors such as prosody [9], referentiality [13] and individual reading span [20]. Individual variation might be expected where it is linked to memory span and pragmatics, but cross-linguistic variation in the absence of grammatical differences is much harder to deal with if locality is taken to be universal.

One reaction to these findings would be to abandon this very hypothesis and assume that locality principles in processing are subject to a certain degree of parametrization. It is widely acknowledged, however, that cross-linguistic variation in parsing principles is highly problematic. Children need to parse the language they hear in order to acquire the grammar of their mother tongue. This will be very hard, if not impossible, if principles of parsing have to be acquired
themselves. And principles of parsing can hardly be acquired as long as there is no grammar to base this process on. As Fodor [9, p. 285] puts it: 'The whole explanatory project [. . . based on the hypothesis that the processing mechanism is fully innate and applies differently to different languages only to the extent that their grammars differ . . . ] is in peril because of the discovery that Late Closure is not universal.' Several accounts for these problematic results have been proposed, with the only agreement being that none of them is completely satisfactory. One position has been to limit the application of structural parsing principles like Late Closure to *primary relations* (e.g. verb-argument relations). Non-primary relations (e.g. modification by RCs) are parsed using a variety of non-structural principles [13]. Other approaches suggest that observed differences in attachment might originate in (i) differences in frequency of exposure to HA vs. LA structures [18]; (ii) parametrization of Predicate Proximity, a principle that interacts with locality [12], (iii) cross-linguistics differences in prosody [9]; (iv) cross-linguistic differences in the relativizing element [15].

An important recent development in this debate came with the discovery, by the applicant [14], of a grammatical confound in the previous literature. Some languages allow an additional grammatical structure, known as Pseudo Relative (PR) in the relevant sentences (2), and there is a large correspondence between PR availability and apparent exceptions to Locality. Pseudo Relatives and RCs, despite being string identical, have very different structural and interpretive properties. Crucially HA is obligatory with PRs, which roughly correspond to English Small Clauses (SC) of the Acc-ing type [4]. The grammatical and interpretive distinction between RCs and PRs clarifies the epiphenomenal nature of the exceptions to Locality and, lifting the burden of explanation from the parser, paves the way for a promising solution to the crisis we discussed.

(2) Vejo o menino que chora  
PR: [VP Vejo [SC o menino que chora]] = I see the boy crying  
RC: [VP Vejo [DP o [NP menino [CP que chora]]]] = I see the boy that cries

2.2. Research plan and methods. In Grillo and Costa [14], I hypothesized that cross-linguistic variation in RC-attachment can be reduced to PR availability, as PRs are asymmetrically available across HA and LA languages. Additionally, I proposed that the exceptional behaviour of HA-languages originates from a parsing preference favouring PRs over RCs (PR-first Hypothesis). This preference is justified on both syntactic and pragmatic grounds as PRs, being a type of Small Clause, are both structurally and interpretively simpler than RCs. Support for these claims comes from both the reconsideration of previous findings (as attachment preferences and PR-availability correspond to a striking degree, both across
languages and structures) and novel results from experiments in which we manipulated PR-availability (Grillo et al 2012a,b, Grillo et al. 2013a,b). In summary, when PRs/SCs are available we find HA, and when only RC interpretation is available we find LA. These are promising findings, yet, further theoretical and experimental work is needed to clarify and strengthen the PR-first Hypothesis. This project proposes 10 studies addressing 3 main issues: (1) cross-linguistic evidence of the interaction of PR-availability and attachment preference, (2) online evidence for PR-first and (3) Discourse vs. syntax in PR-first.

2.2.1. Objective 1. The relation between PR-availability and RC-attachment preference is largely supported by both previous and novel findings. Yet, there are three issues to address. 1. Russian appears to disobey the present generalization and needs to be reassessed. Secondly, our results on Italian and Portuguese are very clear, but these languages have been alternatively classified as HA/LA, making the argument less compelling. Testing Dutch and Spanish, i.e. languages that were consistently classified as HA, will significantly strengthen the proposal. Finally, the status of PR-availability in some languages needs to be clarified (e.g. Japanese). I will design and conduct five versions (Japanese, Russian, Dutch, Spanish, French) of a questionnaire on RC-attachment preferences (Experiment 1.1-1.4). PR-first predicts a strong effect of PR-availability on attachment: i. Predictions for Japanese depend on the PR-status of this language, which is to be studied. ii. Russian, overall LA preference (Hypothesis 1.1). Spanish/Dutch/French, HA in PR-conditions and LA in RC-only conditions (Hypothesis 1.2).  

2.2.2. Objective 2. One prediction of the PR-first Hypothesis is that version [A] of sentence (3) should be harder to process than version [B]. Encountering PR-type verbs like see in [A], the parser initially postulates a PR interpretation of the embedded clause. Further, the subject-verb agreement of the embedded clause forces LA. As HA is obligatory in PRs, the forced LA forces the parser to revise this initial preference and reanalyze the embedded clause as an RC. This reanalysis should cause a cognitive bottleneck (a garden-path effect), which should significantly raise complexity and RTs at the point of disambiguation. Using a stative verb in the matrix clause [B] will force an RC interpretation, since RC attachment is governed by locality, shorter RTs would be predicted (Hypothesis 2.1). The opposite pattern should emerge when HA is forced (Hypothesis 2.2).  

(3) O Eduardo [A]ouviu/[B] vive com os irmos do jovem que estava a cantar no largo. Eduardo heard/lives with the brothers of the youngster (that were) singing in the street. Previous research used stimuli like (3) to test cross-linguistic RC attachment using eye-tracking and ERP (Evoked Response Potentials) and found different results [16,3]. These works, however, didn’t consider PR-availability. We will replicate these studies in Spanish and
Portuguese but explicitly manipulate PR-availability through the main verb as indicated in (3). We predict longer RTs for LA and a P600 effects, i.e. a marker of reanalysis, in the context of PR verbs, and for HA with stative verbs (Hypothesis 2.1, 2.2).

2.2.3. **Objective 3.** The Reference Theory perspective and Interactive models of sentence processing assume that discourse/pragmatics have an early effect on processing. Crain and Steedman [5], Altmann and Steedman [1], for example, demonstrated that processing of RCs is facilitated in the presence of a contextually appropriate discourse representation. In contrast, neither [21] nor [8] found any effect of context in RC-attachment. Using LA-biasing contexts to introduce sentences like those of (3), they found RTs consistent with a generalized HA preference, i.e. discourse information did not appear to influence early stages of RC-attachment. Experiment 3: Once again, this work did not consider PR-availability. We will replicate the study in [8] assuring the context and sentences are RC but not PR compatible, to test for discourse effects on RC-attachment online. Interactive approaches, but not syntax-first approaches, predict early effect of context on RTs (Hypothesis 3.1, 3.2).

2.3. **Expected outcomes / Impact.** The highly relevant blend of knowledge and skills I will acquire through this fellowship will have a long term positive impact in my career. These include:

1. **Broadening of Theoretical Knowledge.** Further investigating the syntactic and pragmatic properties of PRs and RCs, while developing a systematic semantic and syntactic comparison of PRs in Romance and Germanic languages with Internally Headed Relative Clauses Japanese and Korean, which appear to share important properties with PRs, will strengthen my future career in two ways: on the one side, I will further diversify my knowledge in these important domains in theoretical linguistics; on the other, the proposed comparison is both novel and promising as the two structures, despite presenting a number of striking similarities, have been studied in isolation from each other and using different approaches.

2. **Experimental Methods.** The technical competencies I will acquire will perfectly add to my current skills in experimental linguistics and neurolinguistics. Experimental methods such as eye-tracking and ERP are quickly becoming not only essential research instruments in the psycho and neurolinguist toolbox, but are increasingly used in the investigation in theoretical linguistics, which is heading towards the integration of traditional work based primarily on introspection and comparative studies, with sophisticated experimental techniques.

2.4. **3. Managerial, communication and fund-raising skills.** Implementing the project within the excellent academic environment of CLUNL
at the Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa will strongly benefit my career, allowing me to strengthen my managerial, fundraising and writing skills. These skills, which I have been sharpening during the past 4 years at UNL as Ciência Researcher and PI of a FCT project and thanks to the collaboration of outstanding colleagues within the department, are an essential part of the modern scientific work and strategic to obtain research funds that will allow me to establish my own research group and strengthen the capacity for psycholinguistics research at UNL. The Department will support further grant applications both as a follow up and in order to expand the project while its still running. I plan to apply for international research funds which will allow acquiring both an eye-tracker and a EEG scanner. Both technologies are ideal for our department, as they are extremely affordable (especially in comparison to other brain imaging equipments) and, once acquired, have very small running costs.

The project I have designed will strongly help me in establishing and supporting a long lasting research agenda, fully integrated within the host institution’s aim to advance in the path of integrating theoretical and experimental linguistics. The ideas presented here, and the richness of preliminary experimental results that support them, have been received with enthusiasm by the linguistic and psycholinguistic community. This preliminary work, which is undergoing the second round of review in the prestigious international Journal Cognition, was presented at the most important venues in the field of psycholinguistics, including the 25th CUNY Conference on Human Sentence Processing in 2012 (plenary talk) and 2013, and AMLaP 2012 and 2013. This success is due in part to the novelty and importance of the initial discovery, which has the potential to solve many open issues in the theory, but also because it has show a great potential to generate new research. The promising state of this research will open many opportunities for continuing my work in academia. The ability to contribute with groundbreaking research to my field will make my profile highly desirable and will increase my chances of obtaining a permanent research position.

The expertise I will acquire is also very diversified, as it involves formal linguistics, as well as behavioural and neuroimaging methods in psycho and neurolinguistics. The combination of these skills plays an increasingly crucial role in the field and has a demonstrated potential for leading to important results. Strengthening my skills in these areas will allow me to work in strategic yet understudied areas of linguistics.
Given the nature of the project, involving comparisons across a number of different languages, experiments will be conducted at various universities, including Universidade Nova de Lisboa, University College London, Cornell University, CNRS Paris Diderot, University of St. Petersburg and McGill University.

2.5. **Major references.**

3. Career objectives

3.1. Long Term. As PI of an FCT research project at the interface between theoretical and experimental linguistics, I started up and managed the first Psycholinguistics Lab at UNL, where I train and supervise both MA and PhD students in a number of experimental techniques and theoretical issues. This was a defining experience, which strengthened my determination to establish a small, but internationally strong, research group working on theoretical and experimental Psycho and Neurolinguistics. With this goal in mind, I designed the proposed project to strengthen my position in the field, attract international funds and complement and diversify my current knowledge, my methodological expertise and my managerial skills.

3.2. Short Term.

– Goal 1. Eye-tracking experimental methodology: I will learn how to properly design, code and run experiments using eye-tracking technology and how to analyze the results obtained with this technique. Eye-tracking has become standard in current psycholinguistic research
and acquiring it will increase the possibilities for independent and collaborative experimental research.

– Goal 2. ERP - Evoked Response Potentials: I will learn how to design, code, run and analyze experiments using ERP. ERP has been used successfully to investigate a number of essential questions on the relation between brain and behaviour and its role has become increasingly prominent to decide between alternative approaches to sentence processing and representation. This is due to a combination of its remarkable temporal resolution and affordability. As I taught several courses in Neurolinguistics, I already possess a thorough background knowledge about the technique, acquiring this ability will allow me to fully implement this knowledge in my own research.

– Goal 3. Obtain international research funds to purchase a eye-tracker and EEG scanner for the Psycholinguistics Lab at UNL.

3.3. Development / consolidation of an independent career. During the past years I have successfully positioned myself at the intersection of theoretical and experimental linguistics. My previous work on Locality and Agrammatism has had a very influential role in the debate on language development. My most recent work, which forms the basis for this project, has the potential to solve a 30 year old crisis within psycholinguistics. While my collaborators and I have already produced substantial preliminary support for these ideas, the project is both highly innovative and has the potential of being groundbreaking and have a long lasting impact. Having the opportunity to position myself as the first contributor to this line of research will surely have a huge impact on my career development and will open up many opportunities to continue contributing to the fields of theoretical and experimental linguistics. This groundbreaking research has also contributed in further strengthening the international visibility of my home department, the CLUNL. The FCT has played a crucial role in these developments, as the applicant’s research was supported both through the Cincia08 and a FCT Research Grant (PTDC/CLE-LIN/114212/2009) for which I currently am PI. The Grant allowed me to start up the first Psycholinguistics Lab at CLUNL, where most of the preliminary work for this project was conducted in collaboration with other members of the department (particularly Prof. Joo Costa and Prof. Maria Lobo). The host institution offers cutting-edge knowledge in theoretical linguistics and language acquisition and specialized facilities for experimental research in psycholinguistics which perfectly match the profile of the applicant and the goals of the project. Similarly, the project fits perfectly with the Strategic Plan of the Host, which in the past decade has positioned itself as an internationally recognized center of research interfacing theoretical and experimental linguistics. Networking / Internationalization plan Having studied and
worked in 6 countries in Europe and North America (Italy, The Netherlands, Spain, Portugal, UK, Canada). I exposed myself to different and often complementing styles of research and I was successful at developing a wide network of collaborations that at present involves projects with a number of internationally recognized scholars in the field whose collaboration is essential for the proposed project. Theoretical and experimental work in the relevant languages is indicated in parentheses.

– Prof. Ad Neeleman, University College London (Dutch)
– Prof. John Hale, Cornell University (computational models of PR parsing)
– Prof. Junko Shimoyama, McGill University (Japanese)
– Dr. Andrea Santi and Dr. Tuomainen, University College London, Dr. Linnaea Stockall, Queen Mary (ERP, eye-tracking)
– Dr. Barbara Hemforth, CNRS (French, German)
– Dr. Natalia Slioussar, St. Petersburg (Russian)
– Dr. Giorgos Spathas, Stuttgart (Greek)

4. C. Ethical and legal issues

Does your research project / application needs to disclose any statement regarding ethical or legal issues?
Yes

Nature of the ethical/legal issues
I will test my hypotheses using eye-tracking and ERP. I will not test any special population therefore I will recruit the subjects from the student population at university. I include drafts of information sheets for participants and informed consent that participants will be asked to sign. The form addresses purpose of research, methodology, risks/benefits for participants, confidentiality, rights to withdraw from experiment, contact of people involved in the experiment.

5. D. Host Institution


5.2. D2. Description of the host conditions. I will be a member of the Centro de Lingustica da Universidade Nova de Lisboa, at the Faculdade de Cincias Sociais e Humanas da UNL. As a full member of the faculty, I will have access to all the facilities offered by the Division. This includes an office space, the necessary equipment and IT assistance, access to the experimental hardware and software and to the Psycholinguistics Lab I have founded and directed during the past 2 years. Additionally, I will be able to take on supervision and mentoring of Master Students, which I will fully
integrate in the project and train in experimental design, data collection and analysis. I already have experience in training two MAs and two PhD students at UNL.

The applicant has an ongoing collaboration with researchers who have access to native speakers of the languages to be tested and are themselves native speakers of those languages, which further guarantees success of the project. Each of them has already agreed to take part in the present project. Prof. Shimoyama (McGill University) for Japanese, Prof. Neeleman (University College London) and Dr. Korneef (Utrecht institute of Linguistics OTS) for Dutch, Dr. Slioussar (Utrecht and St. Petersburg Universities) for Russian, Dr. Hemforth (CNRS) for French (Objective 1). Access to eye-trackers and EEG scanners is ensured by collaborators at UCL (Dr. Tuomainen and Dr. Santi).