



### Short Visit Grant

### Scientific Report

**Proposal Title:** Learnability traces in word structure: a computational and experimental approach

**Application Reference N°:** 6938

#### 1) Purpose of the visit

My visit aimed to produce a common ground with Dr. Mónica Tamariz in the direction of testing the relative (semantic and syntactic) informativeness of sublexical phonological sequences (like syllables or consonant clusters) beyond traditional morphology. Words that have similar consonants and stress patterns are also similar in terms of their co-occurrence statistics - indicating that consonants underlie a form-function isomorphism which could be useful in learning and/or comprehension. On the other hand, words that share stressed vowels are more different than expected in terms of their co-occurrence statistics, suggesting that these vowels help disambiguation among otherwise similar word forms. Besides, there is a positional asymmetry in terms of phonological information: single phones on the left of a word are better for disambiguation than those on the right (Tamariz 2008).

While extensive work in a few languages has been carried out by Dr. Tamariz and others in this direction (Tamariz 2008, Piantadosi et al. 2009, Farmer et al. 2006), our plan was to combine her experimental setup with a maximally diverse set of languages. This correspond to a

broader conception on how regular properties in languages should be interpreted: hypothetical mechanisms and processes acting upon language structures have to feature in careful experimental setups -thus showing their psychological plausibility- and they have to be attested in a sufficiently large number of languages with independent genealogies – providing in this way a manner of distinguishing between potential and actual influences on language.

Tamariz, Monica. "Exploring systematicity between phonological and context-cooccurrence representations of the mental lexicon." *The Mental Lexicon* 3, no. 2 (2008): 259-278.

Piantadosi, Steven T., Harry Tily, and Edward Gibson. "The communicative lexicon hypothesis." In *The 31st Annual Meeting of the Cognitive Science Society (CogSci09)*, pp. 2582-2587. 2009.

Farmer, Thomas A., Morten H. Christiansen, and Padraic Monaghan. "Phonological typicality influences on-line sentence comprehension." *Proceedings of the National Academy of Sciences* 103, no. 32 (2006): 12203-12208.

## **2) Description of the work carried out during the visit**

A considerable fraction of the time I spent at the University of Edinburgh went to the process of becoming familiar with the experimental paradigms they embrace, in particular the Iterated Learning model (Kirby and Hurford, 2002). The fundamental idea behind this experimental approach to cultural evolution is that learning across generations of agents -for instance, speakers of a language- is driven (among other things) by a pressure towards eased transmission. This is coherent with Dr. Tamariz' and my view on the project, and more in general with the usage-based tradition in linguistics (Ibbotson, 2013).

While originally we thought the setup and the execution of the project was straightforward, a number of considerations of both practical and theoretical nature emerged – in the end, we believe this will result in an improved experiment.

On the testing material, Dr. Tamariz and I came across a series of difficulties related to the task of comparing such effects in languages, and we isolated a number of typological variables (like head-first or dependent-first) that need to be taken into account in order to guarantee

commensurability. This is not necessarily a problem, but it does render the production of the testing material slower.

Among the major theoretical points we discussed at length there was the role of the asymmetry in the relevance of phonological material within words according to their relative position. There exist well documented regularities in the nature of the syntactic and semantic role of morphemes according to their occurrence in word-initial and word-final positions (Bybee et al. 1990). We found in the work of Prof. Michael Ramscar (see for instance, Ramscar et al. 2010) the link between such asymmetry and the learnability component based on contemporary takes on associative learning. The plan is to adopt that framework, which is equipped with a quantitative model -the Rescorla-Wagner equations (Shepard and Allan, 1996). Basically, this neo-associative perspective on learning suggests that encompassing categories (either morphosyntactic or semantic) appearing first in a lexical compound prime the detection of its members, whereas concrete instances appearing before those categories facilitate their acquisition in first place (Ramscar 2013).

In sum, we converged to the following hypothesis: if phonological regularities follow others at the morphosyntactic and semantic level and this finds a functional explanation in a facilitated learning of specific lexical referents, then we would expect (because of the work cited in the previous paragraph) those regularities to occur on word-initial positions. This is exactly what Dr. Tamariz found for Spanish and English so far.

Kirby, Simon, and James R. Hurford. "The emergence of linguistic structure: An overview of the iterated learning model." In *Simulating the evolution of language*, pp. 121-147. Springer London, 2002.

Ibbotson, Paul. "The scope of usage-based theory." *Frontiers in psychology* 4 (2013).

Bybee, Joan L., William Pagliuca, and Revere D. Perkins. "On the asymmetries in the affixation of grammatical material." *Studies in typology and diachrony* (1990): 1-42.

Ramscar, Michael, Daniel Yarlett, Melody Dye, Katie Denny, and Kirsten Thorpe. "The Effects of Feature-Label-Order and Their Implications for Symbolic Learning." *Cognitive Science* 34, no. 6 (2010): 909-957.

Siegel, Shepard, and Lorraine G. Allan. "The widespread influence of the Rescorla-Wagner model." *Psychonomic Bulletin & Review* 3, no. 3 (1996): 314-321.

Ramscar, Michael. "Suffixing, prefixing, and the functional order of regularities in meaningful strings." *Psihologija* 46, no. 4 (2013): 377-396.

**3) Description of the main results obtained**

*(Work in progress)*

**4) Future collaboration with host institution**

Plans for experimental work within the facilities of the University of Edinburgh has been discussed. Furthermore, one of the institutions I am affiliated with -the Linguistics Department of the Max Planck Institute for Evolutionary Anthropology- is analyzing a request of funding for the next stage of our project.

**5) Projected publications / articles resulting or to result from the grant *(ESF must be acknowledged in publications resulting from the grantee's work in relation with the grant)***

*(Work in progress)*

**6) Other comments (if any)**