Short Visit Grant ☒ **or** **Exchange Visit Grant** ☐

(please tick the relevant box)

Scientific Report

Scientific report (one single document in WORD or PDF file) should be submitted online <u>within one month of the event</u>. It should not exceed eight A4 pages.

<u>*Proposal Title:*</u> Lexical acquisition in bilingual contexts: aspects of (extra)linguistic and psycholinguistic modeling

<u>*Application Reference N°*</u>*: 5590*

1) Purpose of the visit

The visit was part of a more general project proposal, in which four applicants have been involved - Marcello Ferro and myself from the Italian National Research Council (Institute for Computational Linguistics), Walter Daelemans and Emmanuel Keuleers, from University of Antwerp and Ghent University, respectively.

The overall research goals of this bilateral action have been to focus on cognitive and extra-linguistic factors involved in bilingual word recognition, with a view to exploring implications, requirements and constraints on computational models of bilingual acquisition. In particular, the cross visits intended to investigate linguistic, extra-linguistic and psycholinguistic issues which are generally assumed to have a bearing on aspects of bilingual lexicon acquisition and word processing, with a view to better understanding their role and dynamic relationship with aspects more closely related to the language-specific input.

In detail, the purpose of my visit was to define the theoretical frame for word similarity and lexical neighbourhood competition, and to explore how computational simulations by using TSOMs can potentially be applied on psycholinguistic issues – e.g. first language acquisition and processing, second language acquisition, input encoding and recoding, relations between fully inflected word forms, and the like.

2) Description of the work carried out during the visit

During the work with Emmanul Keuleers, we focussed on how TSOMs (Temporal Self-Organising Maps, Koutnik, 2007; Ferro et al., 2010; Pirrelli et al., 2011; Ferro et al. 2011; Marzi et al. 2012a, 2012b, 2012c) can provide a rigorous and testable conceptual framework within which to analyse hypotheses for lexical acquisition and processing in a bilingual context, specifically in the context of a second language (L2) acquisition.

In a TSOM, acquisition strategies are analysed by focussing on emergent relations between stored word forms and on dynamic expectation/competition of incoming input. In good accord with a memory-based perspective, these strategies are highly dependent on input properties such as type and token frequency, and semantic and phonological consistency (Lieven & Tomasello, 2008; Tomasello, 2003).

In the proposed unsupervised artificial neural network, the developmental course of word memory traces and their organisation and role in word perception, access and productivity, have been simulated. Since through repeated exposure to time series sequences, nodes become specialised for both nature and context, the temporal connections between consecutively-activated nodes can define the map's expectation for an incoming stimulus. In this perspective, in a bilingual context, many language-specific and extra-linguistics factors affect possible interaction between L1 expectations of incoming stimuli and L2 newly input words to be recognised/recoded and memorised/parsed.

3) Description of the main results obtained

The joint work carried out since the first meeting in Pisa (10th-14th June 2013) has been conducive to an overall assessment of impact and implications of computational simulative frameworks on understanding human behavioural evidence for word processing, access and acquisition.

As already showed in previous work (Marzi et al. 2012a), differential dynamics in L1/L2 acquisition can computationally been simulated by running incremental learning under different training conditions in terms of language-pairs, age and time of L1/L2 exposure. Simulations showed non trivial evolutionary patterns in terms of accuracy in activation and recall on both acquired and novel words in L1/L2, number of processing nodes recruited for the two languages, average error rate of acquired inter-node connections, plasticity of the maps in developing novel connections or parasitically exploiting existing ones.

More results have been running, and they are going to be evaluated from a strictly psycholinguistic plausibility point of view, with the main goal to highlight how prediction and competition based on word similarity and lexical redundancy affect speakers'

anticipation of incoming stimuli, so as to speed input recognition and improve lexical decision. In this perspective, expectation based on L1 affects the way L2 inputs are perceived.

4) Future collaboration with host institution (if applicable)

The joint work started at the first meeting in Pisa (10th-14th June 2013) is going to continue in the coming future with both general and more specific goals in mind.

The acquisition of language-specific phonotactic/orthotactic constraints will be tested by contrasting a parametrically balanced TSOM (i.e. spatio-temporal one), where the spatial contribution tends to categorise instances of the same symbol across different positions in the input, and a time-biased TSOM where the temporal contribution leads to recoding symbols positionally and contextually.

In addition, since the symbol coding we are using for training TSOMs is orthogonal and based on orthographic input, we planned to develop a phonotactic code by using vectors with articulatory features.

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)

In a medium-term perspective, a joint publication will be submitted to an A-Journal in the field of psycholinguistics or mainly focussed on the mental lexicon.

In the short-term, a paper with preliminary results will be submitted to *Suvremena lingvistika*, which will have a special issue dedicated to selected papers based on the contributions presented during the third NetWordS workshop in Dubrovnik (19th-20th September 2013).

In fact, a preliminary study focusing on the definition of the theoretical frame for word similarity and lexical neighbourhood competition, and the application of TSOMs to psycholinguistic issues such as language acquisition and processing, and the emergence of relations between fully inflected word forms, has already been presented during the NetWordS workshop. It has been showed how the entrenchment of patterns of activation, repeatedly and successfully used in word recognition, is key to understanding issues of human word knowledge (Marzi & Ferro 2013).

6) Other comments (if any)

Since a great part of preparatory work has been conducted in advance, the expected duration of the visit in the original proposal has been reduced from 4 to 3

days (from 25.03.2014 to 27.03.2014). The collaborative work with Emmanuel Keuleers has been conducted directly at his affiliation, Center for Reading Research Department of Experimental Psychology, Ghent University, instead of the provisional venue at the University of Antwerp, where the companion short visit (Ferro, 5604) had been hold.

REFERENCES

Bailey T.M., Hahn U. (2001) Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods? *Journal of Memory and Language*, 44, 568-591.

Dijkstra, A., & Van Heuven, W.J.B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175-197.

Ferro M., Marzi C., Pirrelli V. (2011), A Self-Organizing Model of Word Storage and Processing: Implications for Morphology Learning. *Lingue e Linguaggio*, vol. X(2), 209 - 226. Bologna, Il Mulino.

Hernandez, A.E., Li, P., MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, 9(5), 220-225.

Koutnik, J. (2007). Inductive Modelling of Temporal Sequences by Means of Self-organization. In *Proceeding of International Workshop on Inductive Modelling*. Prague. 269-277.

Lieven, E. & M. Tomasello. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. Ellis (eds), *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge. 216-236.

Marzi C. & M. Ferro (2013). Adaptive Strategies in Lexical Acquisition. In *Lingue e Linguaggio,* XII (2), 307-328. Bologna, Il Mulino.

Marzi C., M. Ferro & V. Pirrelli (2012a). Prediction and Generalisation in Word Processing and Storage. In 8th Mediterranean Morphology Meeting Proceedings on Morphology and the architecture of the grammar (MMM8). 113-130.

Marzi C., M. Ferro & V. Pirrelli (2012b). Word alignment and paradigm induction. Lingue e Linguaggio XI, 2. 251-274.

Marzi C., M. Ferro, C. Caudai & V. Pirrelli (2012c). Evaluating Hebbian Self-Organizing Memories for Lexical Representation and Access. In Proceedings of 8thInternational Conference on Language Resources and Evaluation (ELRA - LREC 2012), 886 – 893.

Pirrelli, V., M. Ferro & B. Calderone. (2011). Learning paradigms in time and space. Computational evidence from Romance languages. In M. Goldbach, M.O. Hinzelin, M. Maiden and J.C. Smith (eds*.), Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, 135-157. Oxford: Oxford University Press.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press.