Building the interface between experts and linguists in the detection and characterisation of neology in the field of the neurosciences

Jesús Torres-del-Rey
Facultad de Traducción y
Documentación
Universidad de Salamanca
C/ Francisco de Vitoria 6-16
37008 Salamanca (Spain)
jtorres@usal.es

Nava Maroto
CES Felipe II
Universidad Complutense de
Madrid
C/ Capitán, s/n
28300 Aranjuez (Spain)
mmaro01@ucm.es

1   Introduction

The NeuroNEO Project (García Palacios et al, 2014) aims to collect and suggest new lexical units in close collaboration with specialists in the field and with specialised translators as necessary collaborators and decisive agents in the dissemination of neologisms. Neologisms are a powerful spearhead for science and, needless to say, for the language(s) of science.

Field specialists come in touch with new term creations and concepts both in the process of knowledge reception and knowledge production. In both cases, Spanish specialists' need to conceptualise, verbalise and stabilise the new concepts is often mediated by English as a scientific lingua franca. This results in a diminished capability of non-English speakers to both apprehend and exploit the potential of the new conceptual and linguistic coinages and to cast their own ideas and scientific constructions onto functionally adequate linguistic moulds. We acknowledge that science, at least for the foreseeable future, will be "created" mostly in English, but contend that it can be enriched and expanded by helping it "speak" other languages through translation from and into English.

In the first phase of our project we are concentrating on the knowledge-reception process alone. Our team of Spanish-speaking language, terminology and translation experts and researchers will be monitoring the way neuroscientists cognitively and linguistically cope and deal with conceptual and terminological neologisms when reading recent English articles. To gather this knowledge, we are experimenting with the design of a collaborative tool to collect relevant information of the specialists' encounter with neological occurrences.

2   Defining and detecting neologisms

Neologisms appear in order to fulfil denominative needs. They challenge a language code, pushing it to its limits, but, at the same time, provide the basis for its survival and development (Sánchez Ibáñez, 2013: 58-65). Linguistic classification of neologisms (e.g. morphosyntactic, semantic, loans, calques) would not be beneficial for neuroscientists while processing the scientific information they read, as this would detract from their concentration on the subject matter and "contaminate" the ecology of the activity, by letting an alien expert domain intrude into their habitual knowledge-reception task. We need to provide neuroscientists "ontologically clean" (Winograd and Flores, 1986: 165) linguistic means to help them identify and communicate their neological experiences.

A cognitive-ontological model that has successfully focused on "units of understanding" (rather than static, universal terms or concepts), historical determinations, diachronic evolution and domain expert participation, is Temmerman's sociocognitive approach to terminology (2000), which later led to "termontography" (see, for instance Temmerman and Kerremans, 2003; Kerremans, 2004). Their

main aim is to bring together the activities of terminography and ontology engineering from a socio-cognitive perspective, and to provide a conceptual and methodological basis for the development of applications for different human and semi-automatic terminology-related endeavours.

A "termontology" which defines and describes units of understanding in terms of categories (also specified as human language phrases), intracategorial aspects and intercategorial relationships in a categorisation framework (Kerremans, 2004) seems like an excellent idea if we want to empower domain experts with a familiar map of knowledge which they can contribute to and which can help us to identify systemic/viewpoint change leading to nascent neologisms.

3   Building the interface between field specialists and language experts

The design of the tool must start with the analysis of the activity where it is to be integrated, the concerns and the commitments of the users/actors, and their conceptual model (Liddle, 1996). It is crucial to get as much "practical understanding" (Ehn, 1992: 122-123) of the task that the users will undertake and of potential breakdown; to stimulate the task by not interfering too much in it, unencumbered but enhanced by the interactive and linguistic scaffolding of our tool.

The necessary (hopefully minimal but productive) interference with neuroscientists' cognitive process must be compensated for by satisfying other needs they may (consciously or unconsciously) have or providing them with a rewarding experience (Hassenzahl, 2010). They must be aware that their participation will help science (not only Spanish science) and Spanish as a language of science; at the same time, even those who have no linguistic inclinations must find that, by focusing on the way language represents and creates reality, the tool can help reinforce, deepen and enhance the understanding of the concepts they read and of their potential implications. Ideally, our tool should also become a friendly platform for neuroscientists to read and annotate all their articles.

In order to test what interface, actions, language and workflows could accomplish all this, we will turn to prototyping as a fundamental method in Human-Computer Interaction. "The ability to shape and reshape software requires a capacity for rapid prototyping – for turning an unarticulated idea into a working prototype quickly enough to be able to change it, to listen to it, even to throw it out and to go on to another" (Winograd, 1996: 206).

4   Prototype experiment for expert neologism detection and characterization

The outline of the experiment is as follows: first, our neuroscientist expert subjects will be assigned a recent journal article about some new development in the field in English. The approximate length of each unit of cognitive-textual processing should be one page (a section or a number of sections totalling that length or thereabouts).

As they read, subjects are encouraged to highlight what they consider to be key elements of the text (mono- or plurilexical terms). Afterwards, they will be required to mark concepts or terms that can be considered as new. These new concepts may be expressed through mono- or plurilexical terms, but they can also appear in the form of denominative periphrasis or other kinds of non-lexicalised expressions (García Palacios, 2009: 19; Sánchez Ibáñez, 2013: 295-308). A term or expression will be considered as a neologism if the expert has only been aware of its existence in recent years.

Once the candidate neologisms are identified, experts will have to decide to what broad category of the field of neuroscience the expression identified as new pertains. Then, they will be requested to explain in Spanish and in their own words the novel aspects discovered in the text. The subjects are expected to emphasise in what respect (activities, situations, cases...) new aspects could be relevant.

Finally, experts will have to suggest other terms or categories (both in English and in Spanish) related to the concept they are describing. This last task is meant to detect categories, conceptual frames to which the identified candidate neologism may belong or prototypical characteristics of the concept.

This first pilot experiment should measure the adequacy of the proposal, and in particular:

- The most adequate unit of cognitive processing (section, page, paragraph, time).

- The degree of distortion the experiment imposes on the process of reading, and whether the task implies a cognitive overload.

- To what extent experts are willing to engage in the experiment, whether they find it motivating.

- From a qualitative perspective, which elements experts identify as novel and which not, allowing us to get a feeling of the clues for the identification of new concepts.

5   Tools for the development of the prototype experiment

For the practical development of our prototype experiments we are considering to use the upper categories of the NIFSTD ontology and wikis as a collaborative tool. The availability and suitability for our research of the former has been considered in Maroto (2013).

NIFSTD (NIF Standard) ontology stands out as the most comprehensive ontology of the neurosciences available on the web. Its wide coverage and its degree of normalisation and reusability make this ontology particularly suitable for our research purposes. It is composed of several modules, each one covering a distinct domain of neuroscience (Bug et al, 2008) and reflecting the semantic domains covered by NIFSTD: organism taxonomy, molecules, macroscopic anatomy, sub-cellular anatomy, cell, nervous system function, nervous system dysfunction, phenotypic qualities and investigation.

The upper classes of the ontology will be used in our research as "tags" so that experts can categorise neologisms identified during the reception of specialised articles. Therefore, these upper classes will shape the initial framework of categories from which our top-down strategy starts, which in turn will later be refined thanks to experts' contributions.

At this stage, we need to test to what extent the conceptual classes proposed by existing ontologies can be used as a valuable tool to help experts categorise new terms and concepts. As our research develops, we might want to integrate our findings within an ontology structure (either NIF's or our own).

We also need a collaborative tool for collecting and linking expert knowledge during the experiments. A wiki can be defined as collaborative software that allows different users to develop web sites in a simple way using a simplified mark-up language, based on linking interrelated concepts and tracking changes. A "wiki is a written-down memory with a lot more space than the built-in one, and it's a collective memory, too" (Stafford and Webb 2006). In a wiki "What You Think Is What You Get".

The main reasons that justify the use of wikis in our experiments are that most scientists are already familiar with this kind of software and the fact that wikis are accessible anytime from anywhere with a web connection. Besides, every detail about the editing is stored in the wiki, that is, we can later analyse how much time our experts spent editing, whether they made many or few changes, etc.

Finally, wikis offer the possibility to "tag" pages and to link concepts. Our experts will be required to categorise new terms and/or concepts, that is, to assign one or more cognitive "tags" to the newly discovered concept. Neuroscientists will be given a set of proposed tags from the upper level of NIFSTD ontology to label their responses to our experiments. In this way the suitability of upper-level ontological classes could be tested and, if necessary, the categories could be revised and/or refined.

This contribution presents the theoretical and methodological framework that is the basis of our research. However, a lot of work still lays ahead in order to check the validity of our proposal.

Acknowledgments

Reference

William J. Bug, Giorgio A. Ascoli, Geffrey S. Grethe et al. 2008. The NIFSTD and BIRNLex Vocabularies: Building Comprehensive Ontologies for Neuroscience. *Neuroinformatics.* 6(3), 175-194. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2743139/. [Last access: April 2014].

Pelle Ehn. 1992. Scandinavian Design: On Participation and Skill. In: Paul S. Adler and Terry Winograd. (eds) *Usability: Turning Technologies into Tools*. Oxford University Press, New York, 96-132.

Joaquín García Palacios. 2009. La competencia neológica especializada en el estudio y la actuación sobre la neología terminológica. *Revue Française de Linguistique Appliquée XIV* (2), 17-30.

Joaquín García Palacios, Jesús Torres del Rey, Nava Maroto, Daniel Linder, Goedele De Sterck, and Miguel Sánchez-Ibáñez. 2014. NeuroNEO, una investigación multidisciplinar sobre la neología terminológica. In: Be-

lén Santanta López and Críspulo Travieso Rodríguez (eds). *Puntos de encuentro: los primeros 20 años de la Facultad de Traducción y Documentación de la Universidad de Salamanca.* Ediciones Universidad de Salamanca, Salamanca, 241-260.

Marc Hassenzahl. 2010. *Experience Design: Technology for All the Right Reasons*. Morgan & Claypool. (Kindle Edition).

Koen Kerremans. 2004. Categorisation Frameworks in Termontography. *Linguistica Antverpiensia New Series* 3, 263–277.

David Liddle (An interview with). 1996. Design of the Conceptual Model. Reflective Conversation with Materials. In: Terry Winograd (ed). *Bringing Design to Software*. Addison-Wesley, Reading, MA, 17-131.

Nava Maroto. 2013. Reusing existing conceptual structures and lexica for neology characterization in the field of Neurosciences: the NeuroNEO project. *Proceedings 10th International Conference on Terminology and Artificial Intelligence TIA 2013*: 87-90. Available at: https://lipn.univ-paris13.fr/tia2013/Conference_Proceedings.html. [Last access: April 2014].

Miguel Sánchez Ibáñez. 2013. *Neología y traducción especializada: claves para calibrar la dependencia terminológica español-inglés en el ámbito de la Enfermedad de Alzheimer* (Doctoral thesis). Salamanca: University of Salamanca.

Rita Temmerman. 2000. *Towards New Ways of Terminology Description: The Sociocognitive-Approach*. John Benjamins, Amsterdam/Philadelphia.

Rita Temmerman and Koen Kerremans. 2003. Termontography: Ontology Building and the Sociocognitive Approach to Terminology Description. *Proceedings of CIL17*, Matfyzpress, Prague. Available at: http://taalkunde.ehb.be/sites/www2.ehb.be/files/u96/temmerman_art_prague03.pdf. [Last access: April 2014].

Tom Stafford and Matt Webb. 2006. What is a wiki (and how to use one for your projects). Available at: http://www.oreillynet.com/pub/a/network/2006/07/07/what-is-a-wiki.html?page=1. [Last access: April 2014].

Terry Winograd. 1996. Profile: Hypercard, Director and Visual Basic. In: Terry Winograd (ed). *Bringing Design to Software*. Addison-Wesley, Reading, MA, 206-213.

Terry Winograd and Fernando Flores. 1986. *Understanding Computers and Cognition*. *A New Foundation for Design*. Ablex, Norwood, New Jersey.