Human Language Technologies supporting Therapeutic Practices for Language Disorders: the project STaRS.sys

Gianluca E. Lebani [†‡] & Emanuele Pianta [‡]

[†] CIMeC - Center for Mind/Brain Sciences, University of Trento
Corso Bettini, 31 38068 Rovereto (Trento), Italy
gianluca.lebani@unitn.it

[‡] Fondazione Bruno Kessler
Via Sommarive, 18 38050 Povo (Trento), Italy
pianta@fbk.eu

1. Introduction

A major difficulty in developing technological aids for anomic patients is the need to create tools flexible enough to cope with the great variability of their impairment. As far as therapeutic aids are concerned, the search for flexibility coincides with the need for cognitively motivated models.

In this paper we will introduce STaRS.sys (Semantic Task Rehabilitation Support system), a system for supporting the speech therapist in the preparation of exercises to be submitted to aphasic patients for rehabilitation purposes. We will show how this tool, developed following cognitively plausible statements, is able to support most of the common semantic therapeutic practices for anomic patients. These pages are organized as follows: we begin by briefly introducing naming disorders and therapeutic practices. This will let us identify the system requirements outlined in the third chapter, and the use case scenarios sketched in the remaining of the paper.

2. Background

Aphasia is an acquired language disorder, better thought as a syndrome than a disease, due to a brain damage. The difficulties experienced by aphasic patients can vary substantially, so that different kinds of aphasia have been identified in the literature. Among the aphasia symptoms, the most pervasive and persistent is anomia, that is a difficulty in retrieving words. Anomia itself is not a uniform impairment. Anomic patients can produce very different patterns of naming errors, even if there is a wide consensus in identifying a major opposition between lexical and a semantic impairments.

Therapies for naming disorders too can be classified as either phonological or semantic on the basis of the tasks exploited. We focus on semantic therapy, i.e. on those exercises tapping into the semantic context of a word in order to activate its meaning. The preparation of such tasks often requires the therapist to fill out (by hand) lists of *<concept> feature* pairs like *<apple> is red* and *<nail> has a pointed end*, adopting a representation that has shown to be able to account for several patterns of anomic semantic deficits (cf. [1], [2]).

Features, however, cannot account for the whole variability observed. Other dimensions of variation are word frequency, grammatical class, age of acquisition, grapheme regularity, morphological complexity, abstractness, visual complexity and word length [3]. Typically, the therapist controls for such variables (e.g. concept frequency) by manually checking on the available resources (e.g. a frequency lexicon).

3. STaRS.sys as a CAT Tool

In such a context, STaRS.sys is thought to be used as a helper by a therapist preparing a semantic task, so as a Computer Assisted Therapy (CAT) tool. A challenge in developing similar tools is to design them to be flexible enough to fit the needs of every patient [4]. Such a notion of flexibility is strongly connected to that of cognitive plausibility. That is, the only way for our tool to be useful in a therapeutic context is to be able to cope with the above reported variables that influence the performance of the patients, and it can be achieved only by leaning on a cognitively modeled knowledge base.

According to this statement, we are developing a semantic database in which every concept is associated to the following five kinds of information:
1. Conceptual Taxonomy. Given the importance of categorization in the psychological literature [5], and the existence of category-specific semantic anomias [6], it's vital for our tool to lean on a fully-specified taxonomy, in which every concept belongs to categories like "tools", "living beings" etc.
2. Featural Descriptions. Most authors agree on the central role played by featural descriptions in the semantic memory [5]. Such information can be

exploited for selecting the concepts to be submitted to the patient, e.g. concepts with a specific feature value (e.g. "red objects") or those for which a feature type is particularly relevant (e.g. "animals with a peculiar fur").
3. FEATURE TYPES CLASSIFICATION. A classification of the kinds of features that can be associated to a concept is useful both for selecting feature types of interest or for estimating semantic measures such as feature distinctiveness, semantic relevance, concept similarity and feature correlation ([1], [2]). We proposed and evaluated elsewhere [7] a feature type classification that can be used for such purposes.
4. PROTOTYPICALITY. Concepts can be more or less representative of a category. Controlling for such a variable can influence substantially the outcome of the therapy. Alternatively, working on concepts with different prototypicality can be very informative in highlighting the real nature of the disorder.
5. WORD FREQUENCY. Another well known variable influencing the performance of the patient is the word frequency, seen as an approximation of his/her familiarity with that word [3]. Therefore, this variable is another vital information that our knowledge database must represent for every concept.

4. Use Case Scenario

Built on a lexical infrastructure that provides such semantic information, STaRS.sys can be used for (1) retrieving concepts, (2) retrieving information associated to concepts and (3) comparing concepts. These functionalities will be illustrated by depicting the preparation of three tasks by a therapist (EP) for a patient (gL) with a naming deficit selectively affecting animal concepts.

RETRIEVING CONCEPTS. In a first scenario, the user looks for concepts matching some specifications. By selecting the "Find Concept" modality, the therapist can choose among the following (combinations of) specifications:

- given values for features: e.g. [color = "red"] [1];
- values of prototypicality for given semantic categories: e.g. [semantic category = furniture & prototypicality = "high"];
- values of distinctiveness[2] for given features or feature types: e.g. [color = "red" WITH distinctiveness = "high"] for the feature *is red*;
- values of mean feature distinctiveness[3]: e.g. [mean_feat_dist = "high"];
- values of semantic relevance[4] for given features: e.g. [color = red WITH relevance = "high"];

[1] Illustrative queries are reported in [square brackets]. Two joining operators are used: & when both values refer to the target concept, WITH when one value is a specification of the other.
[2] i.e. the inverse of the number of concepts in which a feature, or a class, appears [8].
[3] i.e. the mean distinctiveness of the whole set of features describing a concept [1].
[4] i.e. a measure of how much a feature distinguishes a concept from other similar ones [2].

- values of frequency: e.g. [frequency = “high”].

The therapist EP exploits this modality for selecting stimuli for a feature generation task, in which patient gL is required to generate phrasal descriptions (i.e. features) for a given set of concepts. EP submits STaRS.sys a query for animal concepts that are frequently used, associated to highly distinctive color features and that have a high mean feature distinctiveness. The output of the system consists of concepts such as *zebra*, *tiger*, *leopard* and *cow*. EP then selects the items of interest to submit to gL for the task.

RETRIEVING INFORMATION ASSOCIATED TO CONCEPTS. By selecting the “Describe Concept” modality, the therapist can choose among the following range and combinations of semantic characteristics:

- feature types: e.g. [feature type = “color”];
- values of frequency, distinctiveness or semantic relevance: e.g. [frequency/distinctiveness/relevance = “high”].

In our fictional therapy, such a modality is useful for preparing a semantic questionnaire, in which gL is asked to mark as true or false a list of concept-description pairs. Our therapist submits the system a query for perceptual or taxonomical highly relevant descriptions of the concept *leopard*. The output of the system consists of short phrases such as *is yellow with black spots* and *is a feline*, that EP pairs to the target concept to prepare the exercise.

COMPARING CONCEPTS. This option is used to find concepts similar to a target concept. EP exploits it to prepare an odd-one-out task, in which gL is required to select the incoherent element out of a triple.

After specifying the reference concept *lion*, EP submits a query for animal concepts that live in a similar/dissimilar habitat. The system outputs a set of similar concepts such as *leopard* and *cheetah*, and a set of dissimilar concepts such as *wolf*. These two sets can be browsed and/or further refined in order to isolate those that are most (or least) similar/dissimilar from the target concept. EP eventually selects a similar and a dissimilar concept that, together with the reference concept, compose the triple that gL will have to analyze.

5. Conclusion and Future Directions

By cross-fertilizing insights from studies belonging to the feature generation paradigm [5] and from techniques developed in the field of common sense knowledge representation (cf. [9]), we are developing STaRS.sys as a CAT tool flexible enough to be used in a therapeutic context.

The usage scenario sketched in these pages illustrates the core skills that our tool must possess. Many other uses and extensions are conceivable. The most straightforward it’s its application to a research context, where many authors have stressed the lack of control for several nuisance variables (e.g.

[1], [6]). Even if some relevant variables are out of the scope of our tool (e.g. visual complexity), the costs of its enhancement would be clearly outperformed by the advantages following the availability of a structured, broadly coverage and systematically accessible resource like STaRS.sys.

Acknowledgments

This work has been realized also thanks to the support from the *Provincia autonoma di Trento* and the *Fondazione Cassa di Risparmio di Trento e Rovereto*. The authors would like to thank Dr. Rita Capasso, Prof. Gabriele Miceli, Dr. Alessia Monti and the CeRiN staff for the useful discussions on the application scenario sketched above.

References

1. Cree, G.S., McRae, K.: Analyzing the Factors Underlying the Structure and Computation of the Meaning of Chipmunk, Cherry, Chisel, Cheese, and Cello (and Many Other Such Concrete Nouns). Journal of Experimental Psychology: General, 132 (2), 163--201 (2003)
2. Sartori, G., Lombardi, L.: Semantic Relevance and Semantic Disorders. Journal of Cognitive Neuroscience, 16 (3), 439--452 (2004)
3. Laiacona, M., Barbarotto, R., Trivelli, C., Capitani, E.: Dissociazioni Semantiche Intercategoriali: Descrizione di una Batteria Standardizzata e Dati Normativi. Archivio di Psicologia, Neurologia e Psichiatria, 54, 209--248 (1993)
4. Petheram, B., Enderby, P.: The Role of Electronic Devices in the Rehabilitation of Language Disorders. In: Stemmer, B., Whitaker, H. (eds.) Handbook of the Neuroscience of Language. Elsevier Science, 427--432 (2008)
5. Murphy, G.: The Big Book of Concepts. MIT Press (2002)
6. Capitani, E., Laiacona, M., Mahon, B.Z., Caramazza, A.: What are the Facts of Semantic Category-Specific Deficits? A Critical Review of the Clinical Evidence. Cognitive Neuropsychology, 20 (3), 213--261 (2003)
7. Lebani, G.E., Pianta, E.: A Feature Type Classification for Therapeutic Purpose: a preliminary evaluation with non-expert speakers. In Proceedings of the 4th ACL-LAW Workshop. 157--161 (2010)
8. Devlin, J.T., Gonnerman, L.M., Anderson E.S., Seidenberg M.S.: Category-Specific Semantic Deficits in Focal and Widespread Damage: A Computational Account. Journal of Cognitive Neuroscience, 10, 77--94 (1998)
9. Lebani, G.E., Pianta, E.: Exploiting Lexical Resources for Therapeutic Purposes: the Case of WordNet and STaRS.sys. In Proceedings of the 2nd COLING-CogALex Workshop. 94--101 (2010)