The UNL Distinctive Features: Inferences from a NL-UNL Enconverting Task

Ronaldo Teixeira Martins,[1] Lúcia Helena Machado Rino,[2]
Maria das Graças Volpe Nunes,[3] Osvaldo Novais Oliveira Jr.[4]

[1]Núcleo Interinstitucional de Lingüística Computacional - NILC
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
ronaldo@nilc.icmc.sc.usp.br

[2]Departamento de Computação - Centro de Ciências Exatas e de Tecnologia - UFSCar
Rod. Washington Luiz, km 235 - Monjolinho - 13565-905 - São Carlos, SP, Brazil
lucia@dc.ufscar.br

[3]Instituto de Ciências Matemáticas e da Computação (ICMC) - Universidade de São Paulo
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
mdgvnune@icmc.sc.usp.br

[4]Instituto de Física de São Carlos (IFSC) - Universidade de São Paulo
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
chu@ifsc.sc.usp.br

Abstract. This paper reports on the distinctive features of the Universal Networking Language (UNL). We claim that although UNL expressions are supposed to be unambiguous, UNL itself is able to convey vagueness and indeterminacy, as it allows for flexibility in enconverting. The use of UNL as a pivot language in interlingua-based MT systems is also addressed.

1 Introduction

Machine Translation (MT) is one of the most controversial subjects in the field of natural language processing. Researchers and developers are often at odds on issues concerning MT systems approaches, methods, strategies, scope, and their potentialities. Dissent has not hindered, however, the establishment of tacit protocols and core beliefs in the area. It has often been claimed that:[1] (1) fully automatic high-quality translation of arbitrary texts is not a realistic goal for the near future; (2) the need of some human intervention in pre-edition of the input text or in post-edition of the output text is mandatory; (3) source language should be rather a sublanguage, and the input text should be domain- and genre-bounded, so that the MT system could cope with natural language ambiguity; (4) the transfer approach is more feasible than the interlingual one, since the latter, albeit more robust and economic, is committed to the

[1] Most of these assumptions can be extracted from the Survey on the State of the Art in Human Language Technology (Cole et al., 1995). Of special interest are the articles concerning multilinguality by Martin Kay (8.1, 8.2) and Christian Boitet (8.3, 8.4).

Universal Network Language: Advances in Theory and Applications.
Research on Computing Science 12, 2005, pp. 27–37.