VoiceUNL: a Semantic Representation of Emotions within Universal Networking Language Formalism Based on a Dialogue Corpus Analysis

Mutsuko Tomokiyo[1], Gérard Chollet[2]

[1] GETA-CLIPS-IMAG & ENST
BP 53 38041 Grenoble cedex 9 France
mutsuko.tomokiyo@imag.fr;tomokiyo@tsi.enst.fr

[2] ENST
46 rue Barrault, 75634, Paris
chollet@tsi.enst.fr

Abstract. The paper aims to propose a semantic representation of emotions for oral dialogues, based on an analysis of real-life conversations, telephone messages and recorded TV programmes, for the purposes of a speech to speech machine translation. Lexicon and phatics are one of important emotion eliciting factors as well as gestures, prosody and voice tone in oral dialogues. So, the semantic representation is made in a way where these factors are taken into account at the same time. Also, it's done within **U**niversal **N**etworking **L**anguage (UNL) formalism, where UW (universal word) plays an important role.

1 Introduction

This work has been carried out in the framework of "VoiceUNL" [21], which is one of subprojects of "LingTour" and "Normalangue"[1] projects. The "VoiceUNL" is an extension of UNL, which is a text-oriented machine translation environment, to oral dialogues.

As for **s**peech to **s**peech **m**achine **t**ranslations (SSMT), the detection of emotions in source languages and its generation in target languages are an important issue from the viewpoint of the naturalness of dialogues [7], because *emotions entails distinctive ways of perceiving and assessing situations, processing information, and prioritising and modulating actions* [24]. It's the key reason for proposing a semantic representation of emotions.

In this paper, we introduce **U**niversal **N**etworking **L**anguage (UNL) briefly in section 2. In section 3, after having surveyed existing approaches to emotion detection, we define emotions and study emotion classes in section 3. In section 4, we detect

[1] The Lingtour and Normalangue projects were launched in 2002 by the partnership which consists of TsingHua University (China), Paris 8 University (France), INT (France), ENST-Paris and Bretagne (France), and CLIPS (France). One of the objectives of the projects resides in R & D to enable multilingual-multimedia MT on user-friendly tools [1].