NESPOLE![1]
Negotiating through Spoken Language in E-commerce

Executive Summary

"E-world fuels U.S. economy" reports USA Today of June 22, citing a Department of Commerce Study, recently unveiled by Vice President Al Gore. The technological leaps in computing and networking infrastructure in the last decade have unquestionably opened up the opportunity of new and exciting applications for the broad general public. Among these, Electronic Commerce (E-commerce) has emerged as the application with the greatest utility for everyday general usage. Up to now, however, E-commerce interactions over the web have mainly been menu based. Current E-commerce technology thus inherently does not support the identification of customer requirements or the any negotiation between the participants. The next generation of E-commerce will need to overcome these limitations. In order to enable E-commerce with human negotiation, the underlying technology must be capable of dealing with human-to-human communication. Furthermore, the reach of electronic markets naturally transcend national boundaries, thus requiring that linguistic boundaries be also removed.

The proposed project - NESPOLE! - will aim to remove the language barriers and enable humans to do business with humans on the internet via an effective cross-lingual human-to-human communication facility. The challenge of such a facility is to provide (1) effective, natural communication between human participants, (2) effective, inobtrusive, and robust language assistance, and (3) easy to use access to multimedia content in a multi-lingual and multi-cultural setting.

Two major scientific obstacles toward this goal will be attacked under this program: (1) the domain limitation of speech translation systems, and (2) the use of advanced multimodal interfaces and their integration with language assistance function in e-commerce, where language assistance must be placed in the context of the (multimedia) content on the internet and evaluated in the context of overall user effectiveness and satisfaction. The former problem of domain limitation is mostly caused by limitations in the machine translation technology that must handle both the ill-formed input from speech and an essentially unlimited world of discourse, making both syntactic and semantic approaches alone difficult. Our approach will be to exploit a mixture of strengths in a multi-engine approach, while also expanding the basic capabilities of these approaches to be robust and portable. The second problem will be addressed by configuring several prototype systems involving actual service providers and by studying multimodal interfaces in the context of extensive e-commerce use.

The project will be carried out by a consortium of six partners in the US and in Europe over a time-frame of three years. Four of the partners (CMU, University of Karlsruhe, IRST, CLIPS) are scientific institutions who have already worked together for many years under the Consortium for Speech Translation Advanced Research (C-STAR) and have an elaborate and solid base of state-of-the art speech translation technology at their disposal, without which such an ambitious project could not be undertaken. Two commercial partners – APT and AETHRA – will provide Web-based and Video-Conferencing infrastructure and a user/customer business community, based on which extensive and realistic user studies will be carried out. Carnegie Mellon's emphasis will be focused predominantly on the machine translation challenges above and (in cooperation with the European partners) on the multimodal interface issues.

[1] *Nespole* (`NES-po-lay`) means literally *loquat fruit* in Italian. As an exclamation, it means *Wow!*.