

# ITALIAN LITERACY TUTOR

## ESPERIMENTI DI RICONOSCIMENTO DI PARLATO INFANTILE

\*Piero Cosi, \*\*Bryan Pellom

\*Istituto di Scienze e Tecnologie della Cognizione - Sezione di Padova "Fonetica e Dialettologia"  
Consiglio Nazionale delle Ricerche, Via G. Anghinoni, 10 - 35121 Padova, Italy

\*\*Center for Spoken Language Research  
University of Colorado at Boulder - Boulder, Colorado, USA

### RIASSUNTO

Vengono descritti gli ultimi sviluppi dell'*ILT – Italian Literacy Tutor* (Cosi et alii, 2004). Nell'ambito di questo progetto, non ancora completamente definito, le ricerche saranno orientate sulla lingua italiana al fine di favorire il miglioramento del successo scolastico da parte degli allievi, mediante l'utilizzo di strumenti educativi tecnologicamente avanzati che dovranno aiutare gli allievi ad imparare a leggere e a comprendere i testi. Per questo progetto saranno utilizzati alcuni strumenti in fase di sviluppo presso *ISTC-SPFD CNR* quali: *LUCIA*, un sistema di animazione facciale *MPEG-4* (Cosi et alii, 2003), *FESTIVAL*, un sistema di sintesi *TTS* (Text-To-Speech) per l'italiano (Cosi et alii, 2001) e *SONIC*, un sistema di riconoscimento automatico di parlato infantile di cui in particolare in questo lavoro vengono descritte le principali caratteristiche di implementazione ed i risultati dei primi esperimenti di riconoscimento.

Il sistema di riconoscimento del parlato *SONIC*, sviluppato dall'Università del Colorado per l'inglese è stato adattato all'italiano utilizzando come materiale vocale il corpus di parlato infantile raccolto presso l'*ITC-IRST* (Giuliani & Gerosa, 2003) nell'ambito del progetto europeo *PF-STAR*. *SONIC* utilizza come front-end il nuovo sistema di analisi denominato *PMDVR* (*Perceptual Minimum Variance Distortionless Response*) (Yapanel and Hansen, 2003) di notevole interesse per le sue ottime prestazioni soprattutto in ambiente rumoroso.

Il tasso di errore di riconoscimento del 13,5% del sistema finale, che comprende la normalizzazione di lunghezza del tratto vocale (*VTLN*), la normalizzazione della varianza dei coefficienti Cepstrali (*CVN*), modelli fonetici addestrati in modalità indipendente dal parlante utilizzando le più recenti strategie iterative denominate *Structural MAP Linear Regression* (*SMAPLR*), risulta in linea con lo stato dell'arte sul riconoscimento del parlato infantile ed è inoltre più che promettente per ulteriori sviluppi futuri. Questi nuovi modelli acustici per il parlato infantile italiano sono stati incorporati all'interno del *CLT* (Colorado Literacy Tutor) quale primo passo verso lo sviluppo della sua corrispondente versione italiana.

### Bibliografia

- P. Cosi, A. Fusaro, G. Tisato (2003). "LUCIA a New Italian Talking-Head Based on a Modified Cohen-Massaro's Labial Coarticulation Model", Proc. Eurospeech 2003, Geneva Switzerland, pp. 127-132, September.
- P. Cosi, R. Delmonte, S. Biscetti, R. Cole, B. Pellom, S. van Vuuren (2004), "Italian Literacy Tutor: tools and technologies for individuals with cognitive disabilities", in InSTIL/ICALL Symposium 2004, Venice, Italy, June.
- P. Cosi, F. Tesser, R. Gretter, and C. A. (with Introduction by Mike Macon) (2001), "Festival speaks italian!" in Proceedings of EUROSPEECH 2001, Aalborg, Denmark, Sept 2001, pp. 509-512.
- B.L. Pellom (2001), "SONIC: The University of Colorado Continuous Speech Recognizer", University of Colorado, tech report #TR-CSLR-2001-01, Boulder, Colorado, March, 2001.
- D. Giuliani and M. Gerosa (2003), "Investigating Recognition of Children's Speech", Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Hong Kong.
- U.H. Yapanel, J.H.L. Hansen (2003), "A New Perspective on Feature Extraction for Robust In-Vehicle Speech Recognition", Proceedings of Eurospeech, Geneva Switzerland 2003, pp.1281-1284