

Linguistically Motivated Text to Speech Technologies for Telecommunication Services with Mixed Language Support

M.Sc. Seminar Abstract

Paulseph-John Farrugia

Supervisor: Mike Rosner

University of Malta*

December 2003

Abstract

Text to Speech (TTS) systems aim to transform *arbitrary* textual input into spoken output by way of the automatic phonetization of the sentences to utter ([8]). Their implementation requires considerable use of both NLP and DSP techniques.

High quality TTS services are one necessary component for improving man-machine communication interfaces and for enabling *Unified Messaging* ([14, 20]), expected to be an important evolution in telecommunications. Briefly stated, the Unified Messaging concept aims to unify different forms of communication media, such as voice, email, fax and instant messaging, providing a single point of access to the user.

To this end, this project aims explores the use of TTS technologies for the purpose of developing a system capable of converting SMS messages to spoken output, thus opening the medium to other formats of delivery. In the local context, SMS messages are generally written in Maltese, English, or, as is quite typical, in a mixture of both languages, a phenomenon arising out of various geographical and sociological reasons ([3]) and often referred to as *code-switching*. SMS messages also generally contain various word contractions and shorthand no-

tation, typical of the SMS lingo. To cater for such issues, stochastic techniques, based upon previous research on language ([2]) and subject ([5]) classification of texts, are proposed.

Additionally, the importance of a framework for describing and assigning *intonation* and other *prosodic* features for more natural TTS results is addressed. Various *acoustic*, *perceptual* and *linguistic* frameworks have been explored for TTS systems ([1, 13, 17]) but there is still no universally accepted formalism. A linguistically motivated annotation system based on *Tones and Break Indices* ([18, 19]) is proposed as a viable formalism for TTS prosody assignment in Maltese, given previous relevant phonological studies ([21]).

The area of TTS for Maltese has been little explored thus far, with only one other successfully implemented system ([15]). It is another objective of this project to aid further linguistic insight into the task, whilst bringing to bear other directly relevant computational linguistic research ([4, 7, 9] amongst others). Advantage is also taken of other available resources ([10, 11, 12, 16]) that support the development of TTS systems and linguistic research.

References

- [1] BAILLY, G., AND BENOIT, C., Eds. *Talking Machines*. North-Holland Elsevier Science Publishers

*This work is supported by MobIsle Communications Ltd. (<http://www.go.com.mt>)

- B.V., Sara Burgerhartstraat 25, P.O. Box 211, 1000 AE Amsterdam, The Netherlands, 1992.
- [2] BEESLEY, K. Language identifier: A computer program for automatic natural-language identification of on-line text. In *Languages at Crossroads: Proceedings of the 29th Annual Conference of the American Translators Association* (12–16 Oct. 1988), pp. 47–54.
- [3] BORG, A. *Ilsienna*. Albert Borg, 1988.
- [4] CALLEJA, S. Maltese speech recognition over mobile telephony. In Cordina and Pace [6].
- [5] CAVNAR, W. B., AND TRENKLE, J. M. N-gram-based text categorization. In *Proceedings of (SDAIR)-94, 3rd Annual Symposium on Document Analysis and Information Retrieval* (Las Vegas, US, 1994), pp. 161–175.
- [6] CORDINA, J., AND PACE, G. J., Eds. *CSAW '03, Computer Science Annual Workshop* (10, 17 July 2003), University of Malta, Computer Science Department.
- [7] DALLI, A. Computational lexicon for Maltese. B.Sc. IT final year dissertation, University of Malta, April 2002.
- [8] DUTOIT, T. *An Introduction to Text-To-Speech Synthesis*, vol. 3 of *Text, Speech and Language Technology*. Kluwer Academic Publishers, P.O. Box 322, 3300 AH Dordrecht, The Netherlands, 1997.
- [9] FARRUGIA, P.-J. An automatic machine translation system for English-Maltese. B.Sc. IT final year dissertation, University of Malta, 1999.
- [10] Festival homepage. <http://www.cstr.ed.ac.uk/projects/festival>.
- [11] Flite homepage. <http://www.speech.cs.cmu.edu/flite/index.html>.
- [12] Freetts homepage. <http://freetts.sourceforge.net>.
- [13] FURUI, S., AND SONDHI, M. M., Eds. *Advances in Speech Signal Processing*. Marcel Dekker, Inc., 270 Madison Avenue, New York, New York 10016, 1991.
- [14] Iec unified messaging tutorial. http://www.iec.org/online/tutorials/unified_mess.
- [15] MICALLEF, P. *A Text To Speech System for Maltese*. PhD thesis, University of Surrey, 1998. unpublished.
- [16] Praat homepage. <http://www.praat.org>.
- [17] SANTEN, J. P. V., SPROAT, R. W., OLIVE, J. P., AND HIRSCHBERG, J., Eds. *Progress in Speech Synthesis*. Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA, 1996.
- [18] SILVERMAN, K., BECKMAN, M., PITRELLI, J., OSTENDORF, M., WIGHTMAN, C., PRICE, P., PIERREHUMBERT, J., AND HIRSCHBERG, J. ToBI: A standard for labeling English prosody. In *ICSLP-92* (1992), vol. 2, pp. 867–870.
- [19] ToBI homepage. <http://www.ling.ohio-state.edu/~tobi>.
- [20] Unified messaging news and information portal. <http://www.unifiedmessaging.com>.
- [21] VELLA, A. *Prosodic Structure and Intonation in Maltese and its Influence on Maltese English*. PhD thesis, University of Edinburgh, 1995.