

Maltese Speech Recognition over Mobile Telephony

Sinclair Calleja

Department of Computer Science and AI,
University of Malta

Automatic speech recognition is the technology that allows humans to communicate to a machine through spoken utterances instead of the more conventional means such as keyboards. The spoken signal is captured and analysed in a series of one or more steps, and action is taken depending on the final representation. Speech recognition inherits much from the cognitive sciences since most approaches try to first understand how humans perceive and interpret speech, and then build a model that best represents human perception.

While models for high performance speech recognition have been studied extensively for many decades, it is only recently that we have seen some impressive results. The main reason for this is that unfeasible models for high performance systems designed in the past have become feasible due to the increase in the available processing power. Modern speech recognisers extract features from the time-varying signal that discriminate between the different units of speech, being affected only slightly by the pitch or manner of the speaker. The features are then passed through one or more trained Hidden Markov Models (HMMs) for the classification of a phonetic sequence, and ultimately the utterance itself [18, 13].

It should be clear that a speech recognition system is highly dependent on the base language. The reasons for this are several, but mainly include (i) the different subset of phones that optimally model a language (ii) the approximately constant language statistics representative of the language, extracted from an annotated speech corpus. All high performance speech recognisers are based on probabilistic models (e.g. HMMs), and thus the importance of (ii) must not be underestimated.

While extensive linguistic study has been performed on the Maltese language in the classical sense, including grammatical studies such as [7], historical studies [3, 8, 10], dialectal studies [2, 9, 1] and phonological studies [6, 4], there currently lacks material for computational resources in Maltese. At the extreme we do not have an organised and central corpus for the Maltese language. Projects requiring such material usually have to consider the task of collecting the different fragments from various sources. With regards to resources related to speech, such as annotated speech corpora, the situation is practically non-existent. Projects in this field [15, 12] have usually relied on developing their own small resource “islands”, limited due to the nature of the laborious task involved in manual annotation. Currently literature includes the development of a speech synthesiser for the Maltese language [15], a tool for automatic annotation of speech samples [12], the development of a computation method for the automatic generation and labelling of a Maltese computational lexicon [11], and projects such as a study in the application of a deterministic spell checker for Maltese [14] and the application of a stochastic spell checker for Maltese [17], and finally an automatic machine translation system from English to Maltese focussing on weather announcements [?]. Various papers have also been published to introduce the ideas, concepts and goals of the Maltilex project, a joint project by the Department of Computer Science & A.I. (Faculty of Science), the Department of Communications and Computer Engineering (Faculty of Engineering) and the Institute of Linguistics, whose major task is the development of a POS-labelled Maltese computational lexicon [19, 20, 16].

In general, the aim of this work is to continue developing on the existing resources, in particular on those concerning speech. A phonetic analysis of the language will be performed from a computa-

tional point of view with the aim of developing and evaluating the first simple speech recogniser for the language. However the project will not stop at the academic level. The intention is to deploy the speech recognition modules on top of an Interactive Voice Response system (IVR), as used extensively by telephony applications (e.g. voice mail). Studies have shown [21] that using a speech interface to IVRs increases successfully serviced customers, reduces customer mistakes, improves routing accuracy, and reduces the routing time. All of these factors will decrease the running costs of an IVR system, in particular those applied to call centres, [22] and provide a much more interesting experience to customers [21]. This naturally yields to a return and eventually profit on the initial investment cost.

From a technical point of view, the GSM mobile channel will be given special consideration due to the fact that GSM encoding is a lossy compression standard that introduced noise to the original waveform [5]. The final deliverable will be a toolkit by which speech-driven IVRs can be easily developed on top of the classical touch-tone system. As a case study a practical IVR implementation will be developed that will focus on the recognition of single words and sequences of continuously spoken digits such as telephone numbers. The actual particular application has yet to be determined since one has to focus the available resources into a commercially viable direction.

The evaluation of the system is expected to provide answers to questions based on [21] such as whether the Maltese public prefers to talk to speech recognition systems in general rather than through touch-tone menus, whether the Maltese public finds the experience of using this particular system better or not than an older touch-tone driven version depending on the application, rate the system on a standard 5-scale Likert-scale and questions of an original nature, such as whether the Maltese public prefers using Maltese or English as a communication medium to IVR systems. An analytical study of the recordings will be performed to determine the extent to which Maltese phonemes are used by Maltese nationals when speaking English words.

Minor tasks include improving the current version of the text-to-phoneme transcriber to extract and publish statistics on Maltese phonemes, both when considered on their own and also in the context of their surrounding phonemes.

References

1. Dionisius A. Agius. Morphological alternatives in the Gozitan dialects of Maltese. In *Materiaux Arabes et Sudarabiques*, pages 111–161. Groupe d’Etudes de Linguistique et de Litterature Arabes et Sudarabiques, MAS-GELLAS, 1992.
2. J. Aquilina and B. S. J. Isserlin. *A Survey of Contemporary Dialectal Maltese*, chapter 4, pages 105–217. University of Leeds, 1981.
3. Joseph Aquilina. Some historical phonetic changes of Maltese. In *Papers in Maltese Linguistics*, pages 117–165. Royal University of Malta, 1961.
4. Marie Azzopardi. *The Phonetics of Maltese: Some Areas Relevant to the Deaf*. PhD thesis, University of Edinburgh, 1981.
5. L. Besacier et al. GSM speech coding and speaker recognition. In *Proceedings of ICASSP’00, Istanbul, Turkey*, 2000.
6. Albert Borg. Elementi per un’ortografia fonemica del maltese. D.Litt. thesis, Università degli Studi, Roma, 1975.
7. Albert Borg. *Ilsienna: Studju Grammatikali*. Published by author, 1988.
8. Alexander Borg. The *Imaala* in Maltese. In *Israel Oriental Studies*, volume 6, pages 191–223. Tel-Aviv University, 1976.
9. Alexander Borg. Reflexes of pausal forms in Maltese rural dialects? In *Israel Oriental Studies*, volume 7, pages 211–225. Tel-Aviv University, 1977.
10. Alexander Borg. *A Historical and Comparative Phonology and Morphology of Maltese*. PhD thesis, The Senate of the Hebrew University, 1978.
11. Angelo Dalli. Computational lexicon for Maltese. Master’s thesis, University of Malta, April 2002.

12. Robert Ellul Micallef. Speech annotation. B.Sc.(Hons) I.T. Dissertation, Department of Communications and Computer Engineering, University of Malta, 2002.
13. B. H. Juang and S. Furui. Automatic recognition and understanding of spoken language—A first step toward natural human-machine communication. In *Proceedings of the IEEE*, volume 88, pages 1142–1165, August 2000.
14. Gordon K. Mangion. Spelling correction for Maltese. B.Sc.(Hons) I.T. Dissertation, University of Malta, June 1999.
15. Paul Micallef. *A text to speech synthesis system for Maltese*. PhD thesis, University of Surrey, 1997.
16. Paul Micallef and Michael Rosner. The development of language resources for Maltese. In *Proceedings of the Workshop on Resources for Minority Languages, Athens, 2000*.
17. Ruth Mizzi. The development of a statistical spell checker for Maltese. B.Sc.(Hons) I.T. Dissertation, University of Malta, June 2000.
18. Lawrence R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. In *Proceedings of the IEEE*, volume 77, February 1989.
19. Michael Rosner et al. Maltilex: A computational lexicon for Maltese. In *Proceedings of the Workshop on Computational Aspects of Semitic Languages, ACL/COLING98, Montreal, 1998*.
20. Michael Rosner et al. Linguistic and computational aspects of Maltilex. In *Proceedings of the ATLAS Symposium, Tunis, May 1999*.
21. Bernhard Suhm et al. A comparative study of speech in the call center: Natural language call routing vs. touch-tone menus. In *CHI 2002: changing the world, changing ourselves*, volume 4, pages 283–290, April 2002.
22. Bernhard Suhm and Pat Peterson. A data-driven methodology for evaluating and optimizing call center IVRs. *International Journal of Speech Technology*, 5:23–37, 2002.