1. Record Nr. UNINA9910831041303321

Autore	Keller Eric
Titolo	Improvements in speech synthesis: cost 258: cost 258: the naturalness of synthetic speech
Pubbl/distr/stampa	[Place of publication not identified], : John Wiley & Sons Incorporated, 2001
ISBN	9786610554690 0-470-84594-5 1-280-55469-X
Edizione	[First edition.]
Descrizione fisica	1 online resource (396 pages)
Disciplina	621.399
Soggetti	Speech synthesis
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	List of Contributors. Preface. PART I: ISSUES IN SIGNAL GENERATION. Towards Greater Naturalness: Future Directions of Research in Speech Synthesis (Keller, E.). Towards More Versatile Signal Generation Systems (Bailly, G). A Parametric Harmonic + Noise Model (Bailly, G.). The COST 258 Signal Generation Test Array (Bailly, G.). Concatenative Text-to-Speech Synthesis Based on Sinusoidal Modelling (Banga, E.R. et al). Shape Invariant Pitch and Time-Scale Modification of Speech Based on a Harmonic Model (O'Brien, D. & Monaghan, A.). Concatenative Speech Synthesis Using SRELP (Rank, E.). PART II: ISSUES IN PROSODY. Prosody in Synthetic Speech: Problems, Solutions and Challenges (Monaghan, A.). State-of-the-Art Summary of European Synthetic Prosody R&D (Monaghan,A.). Modelling F0 Contour in Various Romance Languages: Implementation in Some TTS Systems (Martin, P.). Acoustic Characterisation of the Tonic Syllable in Portuguese (Teixeira, J.P. and Freitas, D.). Prosodic Parameter of Synthetic Czech: Developing Rules for Duration and Intensity (Dohalska, M. et al). MFGI, a Linguistically Motivated Quantitative Model of German Prosody (Mixdorff, H.). Improvements in Modelling the FO Contour for Different Types of Intonation Units in Slovene (Dobnikar, A.). Representing Speech Rhythm (Keller, B.Z. and Keller, E.). Phonetic and Timing Considerations in a Swiss High German TTS System (Siebenhaar, B. et al). Corpus-based

Development of Prosodic Models Across Six Languages (Fackrell, J. et al). Vowel Reduction in German Read Speech (Widera, C.). PART III: ISSUES IN STYLES OF SPEECH. Variability and Speaking Styles in Speech Synthesis (Terken, J.). An Auditory Analysis of the Prosody of Fast and Slow Speech Styles in English, Dutch and German (Monaghan, A.). Automatic Prosody Modelling of Galician and its Application to Spanish (Gonzalo, E.L. et al). Reduction and Assimilatory Processes in Conversational French Speech: Implications for Speech Synthesis (Duez, D.). Acoustic Patterns of Emotions (Pollermann, B.Z. and Archinard, M). The Role of Pitch and Tempo in Spanish Emotional Speech: Towards Concatenative Synthesis (Montero, J.M. et al). Voice Quality and the Synthesis of Affect (Chasaide, A.N. and Gobl, C.). Prosodic Parameters of a 'Fun' Speaking Style(Gustafson, K. and House, D.). Dynamics of the Glottal Source Signal: Implications for Naturalness in Speech Synthesis (Gobl, C. and Chasaide, A.N.). A Nonlinear Rhythmic Components in Various Styles of Speech (Keller, B.Z. ad Keller, Ec.). PART IV: ISSUES IN SEGMENTATION AND MARK-UP. Issues in Segmentation and Mark-UP (Huckvale, M.). The Use and Potential of Extensible Mark-UP (XML) in Speech Generation (Huckvale, M.). Mark-Up for Speech Synthesis: A Review and Some Suggestions (Monaghan, A.). Automatic Analysis of Prosody for Multi-lingual Speech Corpora (Hirst,D.). Automatic Speech Segmentation Based on Alignment with a Text-to-Speech System (Horak, P.). Using the COST 249 Reference Speech Recogniser for Automatic Speech Segmentation (Warakagoda, N.D. and Natvig, J.E.). PART V: FUTURE CHALLENGES. Future Challenges (Keller, E.). Towards Naturalness, or the Challenge of Subjectivenss (Caerlen-Haumont, G.). Synthesis within Multi-Modal Systems (Breen, A.). A Multi-Modal Speech Synthesis Tool Applied to Audio-Visual Prosody (Beskow, J et al). Interface Design for Speech Synthesis Systems (Flach, G.). Index.

Sommario/riassunto

Naturalness in synthetic speech is one of the most intractable problems in information technology today. Although speech synthesis systems have improved considerably over the last 20 years, they rarely sound entirely like human speakers. Why is this so, and what can be done about it? * Prosodic processing must be rendered more varied and more appropriate to the speech situation* Timing, melodic control and the relationships between the various prosodic parameters need increased attention* Signal processing systems must be developed and perfected that are capable of generating more than just one voice from a database* A better understanding must be achieved of what distinguishes one voice from another, and of how speech styles differ between simply reading aloud numbers and sentences and their use in interactive speech * New evaluation methodologies should be developed to provide objective and subjective measurements of the intelligibility of the synthetic speech and the cognitive load imposed upon the listener by impoverished stimuli * Adequate text markup systems must be proposed and tested with multiple languages in real-world situations* Further research is required to integrate speech synthesis systems into larger natural-language processing systems Improvements in Speech Synthesis presents the latest research in the above areas. Contributors include speech synthesis specialists from 16 countries, with experience in the development of systems for 12 European languages. This volume emerges from a four-year European COST project focussed on "The Naturalness of Synthetic Speech", and will be a valuable text for everyone involved in speech synthesis.