



- -

•
•
•



**Kingdom of Saudi Arabi
King Abdulaziz City for Science and Technology
Research Grants Programs**

AT-18-12

FINAL REPORT

An Experimental Arabic Text-To-Speech System

**Dr. Mansour M. Alghamdi
Dr. Moustafa Elshafei Ahmed
Eng. Husni A. Al-Muhtaseb**

King Fahad University for Petrulium And Minarals

2003G

..... ()

..... ()

..... *

..... . .

..... . .

..... . . .

..... . . .

..... . . .

..... . . .

..... . . .

..... . . .

..... . . .

..... *

..... . .

..... . .

..... . . .

..... . . .

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

:

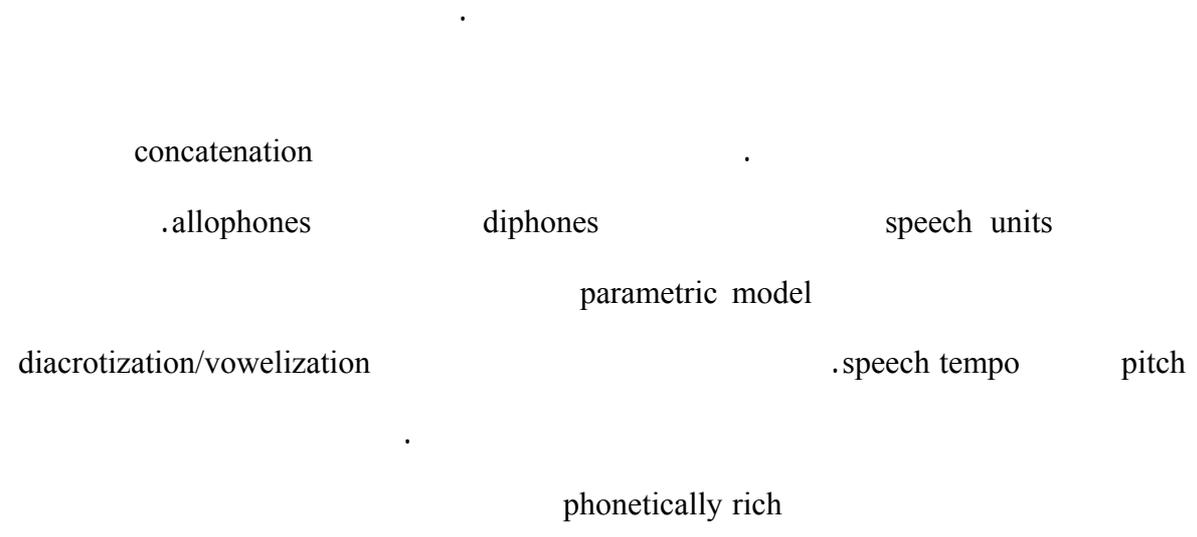
:

:

:

:

Text-To-Speech



Abstract:

In the last 30 years, Text-to-speech methods have been developed for many European languages. However, Arabic text-to-speech (ATTS) has not exceeded laboratories and experiments. This is the first ATTS which is available for researchers and ready for application in different areas such as Human-Machine Communication and reading various Arabic texts. The system is based on a text-to-speech method called concatenation. Two types of speech units were used independently. The first is based on diphone speech units, and the latter is based on the allophone method. A parametric model was built to synthesize the speech and give the user control over the pitch and speech tempo. Since ATTS requires fully vowelized text, while almost all Arabic texts nowadays are unvowelized, a vowelization system was developed. The system is integrated with ATTS. ATTS system has been tested and evaluated. ATTS speech is intelligible, however the system still requires further research and developments in many directions including handling various Arabic texts such as dates, abbreviations, foreign names, numbers and handling prosodic in a natural way.

• • •

• • • • •

• • • • •

• • • • • ()

• • • • • ()

• • • • •

• • • • •

• • • • •

• • • • • -

• • • • •

• •

• • • • •

• • • • •

• • • • • -

• • • • •

• • • • •

• • •

• • • • •

• • • • • -

<http://www.ccse.kfupm.edu.sa/~husni/ICS484/WebPAges/Munawes/Arabic.htm>

- Abu Alyazeed, M. F., M. M. R. Al-Ghoneimy, and M. F. Mohammad. 1989. Comparison of the Syllable and Sub-Syllable Methods for Speech Synthesis. *Proceedings of the Second Conference on Arabic Computational Linguistics, Kuwait.*
- Al-Ani, Salman, H. 1970. *Arabic Phonology - An Acoustical and Physiological Investigation.* Paris. Mouton.
- Al-Dannan, Abdullah and B. Thatouth. 1987. Comprehensive Arabic Morphological Analyzer/Generator. *Technical Report # KSC 029, IBM Kuwait Scientific Center.*
- Alghamdi, Mansour. 1990. *Analysis, Synthesis, and Perception of Voicing in Arabic.* Ph. D. Thesis. Unpublished. University of Reading.
- Al-Jabri, Saad K. 1989. *A Syntactic Analyzer for Arabic.* M. Sc. Thesis, College of Computer Science and Engineering, King Fahd University of Petroleum and Minerals.
- Allen, J., S. Hunnicutt, D. Klatt, 1987. *From Text to Speech: The MITalk System.* Cambridge University Press, Inc.

- Al-Muhtasib, Husni, Moustafa Elshafei and Mansour Al-ghamdi. 2000. Techniques for High Quality Arabic Speech Synthesis. *The Third KFUP Workshop on Information & Computer Science*. 73-82.
- Apple Speech Technologies Home Page. 1998. <<http://www.apple.com/macosp/speech/>>.
- Belhoula, K. 1993. Rule-Based Grapheme-to-Phoneme Conversion of Names. *Proceedings of Eurospeech*. 93 (2): 881-884.
- Bell Laboratories TTS Homepage.1998. <<http://www.bell-labs.com/project/tts/>>.
- Bellcore ORATOR Homepage. 1998. <<http://www.bellcore.com/ORATOR>>.
- Black, A., Taylor P. 1997. Festival Speech Synthesis System: System Documentation (1.1.1). *Human Communication Research Centre Technical Report HCRC/TR-83*.
- Boeffard, O., B. Cherbonnel, F. Emerard, S. White. 1993. Automatic Segmentation and Quality Evaluation of Speech Unit Inventories for Concatenation-Based, Multilingual PSOLA Text-to-Speech Systems. *Proceedings of Eurospeech* (1): 1449-1452.
- Breen, A. 1992. Speech Synthesis Models: A Review. *Electronics & Communication Engineering Journal*. 4: 19-31.
- BT Laboratories Laureate home page 1998. <<http://www.labs.bt.com/innovate/speech/laureate>>
- Campos, G., Gouvea E. 1996. Speech Synthesis Using the CELP Algorithm. *Proceedings of ICSLP*: 3.
- Carlson, R., B. Granström, L. Nord. 1990. Evaluation and Development of the KTH Text-to-Speech System on the Segmental Level. *Proceedings of ICASSP* (1): 317-320.
- Charpentier, F., E. Moulines 1989. Pitch-Synchronous Waveform Processing Techniques for Text-to-Speech Synthesis Using Diphones. *Proceedings of Eurospeech* (2): 13-19.
- Charpentier, F., M. Stella. 1986. Diphone Synthesis Using an Overlap-Add Technique for Speech Waveforms Concatenation. *Proceedings of ICASSP* (3): 2015-2018.
- Coker, C. H. 1985. A Dictionary-Intensive Letter-to-Sound Program. *J. Ac. Soc. Am.*, Suppl. 1, n°78, 1985, S7.
- Dettweiler, H., W. Hess. 1985. Concatenation Rules for Demisyllable Speech Synthesis. *Proceedings of ICASSP 85* (2): 752-755.
- Donovan, R. 1996. *Trainable Speech Synthesis*. PhD. Thesis. Cambridge University Engineering Department, England. <ftp://svr-ftp.eng.cam.ac.uk/pub/reports/donovan_thesis.ps.Z>.
- Dutoit, T. 1996. *An Introduction to Text-To-Speech Synthesis*, Kluwer Academic Publishers. 326 pp.

- Dutoit, T., and H. LEICH. 1993. MBR-PSOLA : Text-To-Speech Synthesis based on an MBE Re-Synthesis of the Segments Database", *Speech Communication*, Elsevier Publisher, November. vol. 13, n°3-4.
- Dutoit, T., V. Pagel, N. Pierret, F. Bataille, O. Vrecken. 1996. The MBROLA Project: Towards a Set of High Quality Speech Synthesizers Free of Use for Non Commercial Purposes. *Proceedings of ICSLP* (3).
- El-Dessouki, A., A. Nazif, O. El-Dessouki, and M. Ahmed. 1987. An Expert System for Understanding Arabic Sentences. *10th National Computer Conference & Exhibition*, King Abdul-Aziz University, Jeddah.
- El-Emam, Y. A. and A. Dannan. 1986. The Phonetics of Modern Standard Arabic. *Technical Report # KSC 022, IBM Kuwait Scientific Center*.
- El-Imam, Y. A. 1988. Synthesis of Arabic Speech from Sub-Syllables. *Proceedings of the Regional Conference on Informatics and Arabization, Tunis*.
- El-Imam, Yousef A. 1989. Unrestricted Vocabulary Arabic Speech Synthesis System," *IEEE Trans. Acoustic, Speech and Signal Processing*. ASSP-37, (12): 1829-1845.
- Elsadany, T. A. and M. A. Hashish. 1987. Semi-Automatic Vowelization of Verbs. *10th National Computer Conference & Exhibition*, King Abdul-Aziz University, Jeddah.
- Elshafei, M. Ahmed and A. Al-Sultan. 1989. A Letter-to-Sound Algorithm for Automatic Conversion of Voweled Arabic Text to Speech. *Proceedings of the 11th National Computer Conference, Dhahran*.
- Elshafei, M. Ahmed. 1989. Allophone-Based Arabic Speech Synthesis", *Proceedings of the 11th National Computer Conference, Dhahran*.
- Elshafei, M. Ahmed. 1991. Toward an Arabic Text-To-Speech System. *The Arabian Journal for Science and Engineering*, Volume 16, Number 4.
- Elshafei, M. Ahmed, M. S. Ahmed, and A. M. Marzooq. 1988. An Arabic Text-to-Speech System. *Proceedings of the Regional Conference on Informatics and Arabization, Tunis*.
- Elshafei, Mustafa, Husni Al-Muhtasib and Mansour Alghamdi. 2002. Techniques for high quality Arabic speech synthesis, *Information Science*. 140 (3-4): 255-267.
- Fant, G. 1970. *Acoustic Theory of Speech Production*. Mouton, The Hague.
- Flanagan, J. 1972. *Speech Analysis, Synthesis, and Perception*. Springer-Verlag, Berlin-Heidelberg-New York.
- Flanagan, J., Rabiner L. (Editors). 1973. *Speech Synthesis*. Dowden, Hutchinson & Ross, Inc., Pennsylvania.
- Ghazali, Salem. 1983. Elements of Arabic Phonetics. *Proc. of the Arab Science and Technology on Applied Arabic Linguistics and Signal and Information Processing*, Rabat, Morocco.

- Hilal, Y. 1989. Arabic Morphological Generator. *Proceedings of the Second Conference on Arabic Computational Linguistics*, Kuwait.
- Holmes, J., I. Mattingly, J. Shearme. 1964. Speech synthesis by rule. *Language and Speech*. Vol 7: 127-143
- Huang, X., Acero A., Adcock J., Hon H., Goldsmith J., Liu J., Plumpe M. 1996. Whistler: A Trainable Text-to-Speech System. *Proceedings of ICSLP96* (4).
- Huang, X., Acero A., Hon H., Ju Y., Liu J., Mederith S., Plumpe M. 1997. Recent Improvements on Microsoft's Trainable Text-to-Speech System - Whistler. *Proceedings of ICASSP97* (2): 959-934.
- Hunnicut, S.1980. Grapheme-to-Phoneme rules : a Review. *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR 2-3*, 38-60.
- Khayat, M. 1989. Directions in Natural Arabic Understanding. *Proceedings of the Second Conference on Arabic Computational Linguistics*, Kuwait.
- Klatt, D. 1987. Review of Text-to-Speech Conversion for English. *Journal of the Acoustical Society of America, JASA* vol. 82 (3), pp.737-793.
- Kleijn, K., Paliwal K. (Editors). 1998. Speech Coding and Synthesis. *Elsevier Science B.V., The Netherlands*.
- Laine, U., Karjalainen M., Altosaar T. 1994. Warped Linear Prediction (WLP) in Speech Synthesis and Audio Processing. *Proceedings of ICASSP94*: 349-352.
- Lisker, L. 1986. Voicing in English a catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech*: 29, 3-11.
- Markel, J. D., and A. H. Gray Jr. 1976. *Linear Prediction of Speech*. Springer Verlag, New York, pp. 10-42.
- Maryati, M. 1985. Speech Processing: Application to the Arabic Language. *Workshop on Computer Processing of the Arabic Language*, Kuwait.
- MBROLA Project Homepage. 1998. <<http://tcts.fpms.ac.be/synthesis/mbrola.html>>.
- McAulay, R., Quatieri T. 1986. Speech Analysis-Synthesis Based on Sinusoidal Representation. *Proceedings of ASSP-34* (4): 744-754.
- Mouradi, A., A. Rajouai, and M. Najim 1985. Unlimited Vocabulary Synthesis for Arabic Language. *Proceedings of the 4th Int. Conf. on Digital Signal Processing in Communications*, Loughborough Univ., England.
- Nassir, Abdulmunim abdulmir. 1985. *Sibawayh the Phonologist: A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in His Treatise Al Kitab*. unpublished Ph. D. Thesis, University of York, England.
- Rabiner, L., Shafer R. 1978. *Digital Processing of Speech Signals*, Prentice-Hall.
- Santen, J., Sproat R., Olive J., Hirschberg J. (editors) 1997. *Progress in Speech Synthesis*, Springer-Verlag New York Inc. (Includes CD-ROM).
- SoftVoice, Inc. Homepage.1997. <<http://www.text2speech.com/>>.

- Spiegel, M. 1993. Using the ORATOR Synthesizer for a Public Reverse-Directory Service: Design, Lessons, and Recommendations. *Proceedings of Eurospeech* (3): 1897-1900.
- SVOX Text-to-Speech Synthesis Homepage. 1997. <<http://www.tik.ee.ethz.ch/cgi-bin/w3svox>>.
- Tatham, M., Lewis E. 1992. Prosodics in a Syllable-Based Text-to-Speech Synthesis System. *Proceedings of ICSLP92* (2): 1179-1182.
- Tatham, M., Lewis E. 1995. Naturalness in a High-Level Synthetic Speech System. *Proceedings of Eurospeech 95* (3): 1815-1818.
- Thabit, Khaled Omar. 1998. Design of Knowledge-based System for Correct Arabic Pronunciation. *The Second KFUPM Workshop on Information & Computer Sciences*. Dhahran. 81-94.
- Valbret, H., Moulines E., Tubach J. 1991. Voice Transformation Using PSOLA Technique. *Proceedings of Eurospeech 91* (1): 345-348.
- Ya'akov Gal. 2002. An HMM Approach to Vowel Restoration in Hebrew and Arabic *ACL 02 Semitic Language Workshop*.