

CONTACT
INFORMATION

Newtownmoynagh, Trim, Co. Meath, Ireland
+353-86-6061517
mtoohier@computing.dcu.ie

EDUCATION

Ph.D. in Computing

Dublin City University, Glasnevin, Dublin 9

pending viva voce

Dissertation title: Machine Learning of Glottal Source Variation

Abstract: In an effort to positively influence the area of automatic speaker characterisation, we draw on existent technologies to efficiently derive the voice source and subsequently learn its behaviour in a variety of environments.

We observe the behaviour of the Liljencrants-Fant (LF) parameters as fitted to closed-phase inverse filtered vowel data from a corpus of two speakers, male and female. The LF parameters are tracked as they vary across vowel, phonetic environment, duration, power and fundamental frequency. Statistical analysis is applied to the data in an effort to pinpoint those phenomena that attribute to variation in parameters within and across speakers. Non-linear regression trees and back-propagation neural networks are built for each parameter and tested for learning capability.

The learnt knowledge is applied to recognition and synthesis experiments in order to gauge the performance of such speech technologies when speaker-specific information is included. We believe that this knowledge of voice source variation both within and among speakers has the potential to influence not only speech recognition and synthesis, but many areas of speech technology.

B.Sc. in Applied Computational Linguistics

Dublin City University, Glasnevin, Dublin 9

November 2001

Degree focus: The B.Sc. in Computational Linguistics was primarily aimed towards Natural Language Processing (NLP) and developing the skills needed for a career in many areas of NLP. Areas studied included Speech Processing, Algorithms and Data Structures, Compilers, C++ Programming, Linguistics, German, Translation, Computer Aided Language Learning (CALL), Machine Translation, Logic, Syntax and Semantics.

A final year project was integral to the degree and my approach was largely an investigative study with subsequent development of a pronunciation tool for German language learners, incorporating both CALL and speech processing disciplines. The end product instructed users how to achieve better German vowel pronunciation using both visual and audio cues. Visually, a graph of the vowel space was provided encouraging users to imitate a target vowel position, based initially on their current first and second formant measurements when uttering a specific vowel and secondly on the changes they needed to make in order to achieve what was computed to be their *target* first and second formant positions based on equations derived from male/female tags and the first three formants of a database of native German speakers. Audio examples were also provided for the users convenience.

**WORK
EXPERIENCE**

Exam Invigilator Exams Office, Dublin City University
2004 - present
Senior Exam Invigilator

Exam Invigilator Exams Office, Dublin City University
2002 - 2004
Standard Exam Invigilator

Phonetician VoiceLogics
September 2002 - February 2003
Employed to provide transcriptions for Irish place and company names in regional Irish accents for use in a recognition system geared specifically towards the Irish market.

Tutor North-Dublin Access Programme, Dublin City University
each September 2002 - 2004
Tutored MS office applications to incoming 1st year students from disadvantaged areas as part of a two week North Dublin Access Program (NDAP).

Tutor School of Computing, Dublin City University
2001-2004
Acted as laboratory tutor to Computational Linguistics, Computer Applications and Masters students. Subjects included Speech Processing, Linguistics, Digital Signal Processing and Biometrics. Duties included marking assignments and projects and acting as stand in lecturer. I also tutored a class of 60 Computer Application students in Logic.

Intern School of Computing, Dublin City University
June - September 2001
Coded and streamlined a Kalman Filter based approach to closed-phase glottal source analysis.

AWARDS

2004 ISCA
Student grant for attendance of ICSLP-Interspeech Conference, Korea (accommodation and registration costs)

2001 Enterprise Ireland
Three year top-up grant during PhD studies, totalling EUR7,620, Ref. No. BR/2001/158

2001 School of Computing, Dublin City University
Funding of EUR50,000 to pursue my Ph.D. research

2001 Dublin City University
14 week Summer Internship working on Speech Processing applications

SKILLS

Technical expertise:

- MATLAB programming
- Speech recognition and synthesis technologies
- Machine learning (Classification and Regression Trees, Artificial Neural Networks, Hidden Markov Models)
- Speech analysis and processing toolkits (Praat, HTK, XWaves, MBROLA), automatic corpora alignment

- Statistics for speech (ANOVA, linear and nonlinear regression, Pearson Chi-square correlation coefficients, descriptive statistics)
- Latex, HTML, Word, Excel, Powerpoint

Extensive use and application of

- Perl, C++
- NLP
- Phonetics

Further skills:

- Fluency in German, beginners French
- Developing and organizing perceptual testing
- Corpora collection, including preparing sound studio, instructing test subjects, development of test material
- Writing funding applications (ISCA and Enterprise Ireland awards received)

PUBLICATIONS

Tooher, M. and McKenna, J.G. 2005 *Application of Glottal Source Information to Speech Recognition*. Submitted for review to EUROSPEECH-Interspeech, Lisbon, 2005.

Tooher, M. and McKenna, J.G. 2004 *Prediction of the Glottal LF parameters using Regression Trees*. In Proceedings of ICSLP-Interspeech, Korea, 2004.

Tooher, M. and McKenna, J.G. 2003 *Variation of the glottal LF parameters across F0, vowels, and phonetic environment*. In Proceedings of VOQUAL Workshop, Geneva, 2003.

CONFERENCE PRESENTATIONS

ICSLP-Interspeech'04, JeJu Island, Korea

VOQUAL Workshop'03, Geneva, Switzerland

(See Publications section for details.)

INVITED PRESENTATIONS

School of Computing, DCU March 2003

Transfer talk: Presentation to School of Computing staff and postgraduates of research achievements and proposed future research on successful entry to the PhD program

National Centre for Language Technology Seminar Series February 2003

Machine Learning of Speaker Characteristics: Research introduction and progress report

School of Computing Postgraduate Seminar Series, DCU May 2002

Machine Learning of Speaker Characteristics: Research introduction

National Centre for Language Technology Seminar Series March 2002

Hidden Markov Models: Tutorial

University of Limerick February 2002

A Kalman Filter based approach to Closed Phase Inverse Filtering: Internship report

REFEREES

Dr. John McKenna

Lecturer

School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland

john.mckenna@computing.dcu.ie

+353-1-7005506

Prof. Josef van Genabith

Associate Professor

School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland

josef.vangenabith@computing.dcu.ie

+353-1-7005074

Dr. Darragh O'Brien

Lecturer

School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland

darragh.obrien@computing.dcu.ie

+353-1-7005645

**RESEARCH
INTERESTS**

During the course of my Ph.D. I had the opportunity to immerse myself in many areas of speech technology, statistical analysis techniques, and machine learning disciplines. As a result of this research, my primary interest lies in further exploring the role of speaker characteristics in improving synthesis and recognition applications.

The results of my research highlighted issues which need further attention. Although the results suggest that up to 80% of voice source variation for a given speaker depends on the prosodic and contextual environment, one aspect which was not accounted for concerns the paralinguistic variation, i.e., dynamic changes in the voice which reflect the speakers state, mood, emotion and attitude (angry, happy friendly, etc.). This aspect is particularly challenging, as it very often involves considerable shifts in voice quality, away from the relatively more modal voice. Application of more independent variables to statistical investigation such as this perceived emotion, current health status, and age, among many other possible variables is also a requirement. The 20% (approximately) of data which the present system does not predict almost certainly reflects this missing dimension, which will be of crucial importance for the future enhancement of synthesis.

Recognition has recently been shown to benefit by inclusion of prosodic information. My Ph.D. research has also displayed improvements not by including prosody directly, rather by the inclusion of speaker characteristic glottal source information as learned from varying prosody. A novel approach would be to combine both approaches and include both prosodic models and adjust the spectrum according to predicted glottal behaviour resulting from prosodic and contextual variation.

I am particularly interested in determining the dynamics and dependencies of characteristics observed; whether speaker characteristics are deep rooted or can they be altered by veering away from ones natural style of speaking and if so can these alterations be tracked and learned.

In order to continue the automatic nature of this research, further attention also needs to be applied to efficient large scale automatic inverse filtering of speech and elements of source filter interaction present in the resulting glottal source waveform. Often problems encountered when investigating glottal source variation were attributable to inverse filtering inadequacies or the presence of source-filter interaction. A more robust preface to glottal source analysis and parameter learning would most certainly improve subsequent application to synthesis and recognition.

I believe that further research in the area of speaker characteristics has much to offer many technologies including speaker identification and verification, speech recognition, speech synthesis, speech pathology and forensics.