



HAL
open science

A Multi-modal System for Public Speaking

Fiona Dermody, Alistair Sutherland, Margaret Farren

► **To cite this version:**

Fiona Dermody, Alistair Sutherland, Margaret Farren. A Multi-modal System for Public Speaking. 15th Human-Computer Interaction (INTERACT), Sep 2015, Bamberg, Germany. pp.499-501, 10.1007/978-3-319-22723-8_47 . hal-01610865

HAL Id: hal-01610865

<https://hal.inria.fr/hal-01610865>

Submitted on 5 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

A Multi-Modal System for Public Speaking

Pilot Study on evaluation of real-time feedback

Fiona Dermody¹ Alistair Sutherland² Margaret Farren³

^{1,3}Faculty of Humanities, ^{1,2}Faculty of Engineering and Computing
Dublin City University,
Glasnevin, Dublin 9.
Ireland.

Fiona.Dermody3@mail.dcu.ie

Abstract.

A prototype has been developed for a digital system and multi-modal user interface to analyze social signals displayed during public speaking. User testing on the prototype has commenced to evaluate the most effective way to display real-time feedback to users on their speaking performance.

Keywords. Affective Computing, Multi-Modal Interfaces, Social Signal Recognition, Human Computer Interaction

1 Introduction

The fear of public speaking tops the list of human phobias. However, success in social, academic and occupational situations depends on the ability to communicate effectively to groups. A fear of public speaking thus limits achievement in social gatherings, education and enterprise [1]. However, the fear of public speaking can be so great that it can lead to avoidance of speaking in the public domain altogether [2]. The problem is a recursive one, how can an individual improve their speaking skills and reduce their fear of speaking in public if they avoid speaking in public?

Public speaking is not just about the words spoken. Effective speaking involves the use of gestures, facial expressions and vocal variety. All these social signals combine to give the appearance of self-confidence in a speaker. Research has found that anxious speakers do not engage the attention of an audience [3]. Rather the audience focuses on the speaker's nervous disposition instead of their words [4].

The solution, which we propose, is interdisciplinary. It incorporates theory from computing, psychology and communications. It is envisaged that this social signal recognition system will enable individuals to develop their competence in public speaking.

Using a combination of 3D video imaging, audio and social signal processing algorithms, this digital system analyses facial expressions, tone of voice and gestures. The system then provides feedback on the user's speaking performance. It will also deliver tutorial videos on good speaking practices. Exposure to these dynamic features will enable a speaker to systematically develop confidence and skill before speaking in front of a live audience.

Skilled human trainers in communication are scarce and expensive. This digital system incorporates experience from one such skilled human trainer to provide constructive feedback to users on their speaking performance. This digital experience will enable anxious speakers to develop their public speaking skills cost-effectively, in private and at their own pace.

1.1 Existing Research

There have been attempts to use social signal recognition for public speaking but not for instruction purposes [5]. This system is innovative because it will extend the field of human computer interaction:

- Combining all modalities - voice, gesture, facial expression and body pose into a multi-modal system for delivering instruction in public speaking
- Incorporating the knowledge of experts in public speaking ensures that the feedback provided on a user's speaking is based on a real-world, practice-based approach

2 Research Question

Which form of real-time multi-modal feedback is most effective for users to develop skill in public speaking?

3 Proposed multi-modal system for public speaking

The system will use a Microsoft Kinect connected to a computer. Social Signal Processing techniques will be used to recognize the speaker's body language, gestures, voice, and facial expressions

- Classify speaker's emotion, as perceived by the audience, from the combination of the above
- Perceived emotion is regarded as the primary component for analysis.
- Give feedback to user on speaking performance

- Provide examples of good and bad speaking practice
- Set tutorial exercises from beginners level to advanced level and will evaluate the user's performance

4 Progress to date

A prototype has been developed following a user survey on the features required in a multimodal system for public speaking. Initial user testing has commenced.

5 Expected Results

One of the technical challenges to be overcome during the development of the system is the optimal way to display feedback to the user in real time during their speaking task. The nature of the feedback is imperative, as our survey showed that some users prefer visual feedback while others prefer textual feedback. The system will be evaluated to ascertain what is the most effective way to display feedback to users on their speaking performance. Results of this user testing will be presented.

Acknowledgement

This material is based upon works supported by Dublin City University under the Daniel O'Hare Research Scholarship scheme.

References

1. Dwyer, K. & Davidson, M. (2012) Is Public Speaking Really More Feared Than Death? In *Communication Research Reports*, 29:2, 99-107
2. Feldman, P., Cohen S. and N. Hamrick, Psychological stress, appraisal, emotion and cardiovascular response in a public speaking task. In *Psychology & Health*, 2004.
3. Freeman, T. and Sawyer, C. Behavioral inhibition and the attribution of public speaking state anxiety. In *Communication Education*, no. March 2012, pp. 37-41, 1997
4. Winton, E., Clark, D. and R. Edelman. Social anxiety, fear of negative evaluation and the detection of negative emotion in others. In *Behaviour Research and Therapy*, vol. 33, no. 2, pp. 193-196, 1995.
5. Batrinca, L.M., Stratou, G., Shapiro, A., Morency, L.-P. and S. Scherer. Cicero - towards a multimodal virtual audience platform for public speaking training. In *Proc. 2013 International Conference on Intelligent Virtual Agents*, pp. 116-128, 2013.