PAPER FOR PATHOLOGY REVIEW

SAILORS, Shrikanth Narayanan

Signal Analysis and Interpretation Lab (SAIL),
University of Southern California, Los Angeles, California, USA

ABSTRACT

***Index Terms*—** Pathological speech, intelligibility

1. INTRODUCTION

2. RELATION TO PREVIOUS WORK

3. CHALLENGES IN THE DOMAIN OF UNDERSTANDING PATHOLOGICAL CONDITIONS

Theodora and Rahul

3.1. Defining pathological speech

3.2. Subjective Impressions

3.3. Speaker variability

3.4. Approaches by speech scientists

3.4.1. Obtaining data

Annotations, data demographics, data collection environment

3.4.2. Features

Previous studies have attempted to capture the wide variability of pathological speech through various acoustic and phonological features, as well as non-verbal discourse markers.

Voice quality and prosodic features have been extensively used because of their high interpretability and computational efficiency [1, 2, 3], while multi-scale spectro-temporal modulation indices attempt to represent the irregular spectral perturbations of pathological speech [4, 5, 6]. Vocal source excitation and articulatory features have been proposed in order to capture the malfunctioning of various parts of the speech production system caused by vocal disorders [5, 7]. Other efforts have focused on developing distance measures between healthy and pathological speech [8]. These frame-level features can be incorporated into long-term measures through phone or utterance level functionals [9], contour parameterization [10], and other non-linear transformations [9, 11, 12].

ASR can yield confidence indices of normal speech through lattice posteriors and recognition accuracy metrics [9, 13, 14, 15, 16]. ASR output is further able to provide durational features at the syllable and word level that can be indicative of atypicality [11, 17]. Despite the knowledge-driven nature of this approach, challenges of using ASR metrics include the potentially limited vocabulary size, the existence of sparse multilingual data, and the need for speaker-dependent acoustic models.

Non-verbal vocalizations are an essential part of spoken communication for regulating and coordinating discourse. Their atypical occurrence and expression has been related to various neurological and mental disorders [18]. Previous studies have examined the role of fillers, pauses, and laughter in pathological speech [19, 18, 20].

The inherently diverse information present in the speech signal, such as speaker traits, gender and age effects, environmental conditions, etc., makes it hard to disentangle actual pathology-dependent conditions from other factors. Although previous studies have indicated strong correlates of many of the aforementioned features to pathological constructs, careful methodological and experimental planning has to be conducted in order to make sure that the segmentation of the acoustic features space is performed in terms of the relevant pathological effects [21]. Towards this direction, ecological data capture procedures, reduced-size interpretable features, appropriate statistical analysis, and legitimate experimental validation are encouraged.

3.4.3. Machine learning

4. CASE STUDIES

Jangwon, Naveen, Danny, Jimmy, Bo ....

1. Pathology speech challenge Jangwon, Naveen
2. Parkinsons challenge Jangwon, Rahul ..
3. Depression work
4. Other work in ASD, addiction etc. Danny, Jimmy, Bo

5. REFERENCES

[1] J.P.H. Van Santen, E.T. Prud'hommeaux, L.M. Black, and M. Mitchell, "Computational prosodic markers for autism," *Autism*, 2010.

[2] A. Tsanas, M. Little, P.E. McSharry, J. Spielman, and L.O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1264–1271, 2012.

[3] D. Bone, C.C. Lee, M.P. Black, M.E. Williams, S. Lee, P. Levitt, and S. Narayanan, "The psychologist as an interlocutor in autism spectrum disorder assessment: Insights from a study of spontaneous prosody," *Journal of Speech, Language, and Hearing Research*, vol. 57, no. 4, pp. 1162–1177, 2014.

[4] J.M. Liss, S. LeGendre, and A.J. Lotto, "Discriminating dysarthria type from envelope modulation spectra," *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 5, pp. 1246–1255, 2010.

[5] T. H. Falk, W.Y. Chan, and F. Shein, "Characterization of atypical vocal source excitation, temporal dynamics and prosody for objective measurement of dysarthric word intelligibility," *Speech Communication*, vol. 54, no. 5, pp. 622–631, 2012.

[6] J.R. Williamson, T.F. Quatieri, B.S. Helfer, J. Perricone, S.S. Ghosh, G. Ciccarelli, and D.D. Mehta, "Automatic recognition of unified Parkinson's disease rating from speech with acoustic, i-vector and phonotactic features," in *Proc. Interspeech*, 2015.

[7] S. Hahm and J. Wang, "Parkinson?s condition estimation using speech acoustic and inversely mapped articulatory data," in *Proc. Interspeech*, 2015.

[8] Lingyun Gu, John G Harris, Rahul Shrivastav, and Christine Sapienza, "Disordered speech assessment using automatic methods based on quantitative measures," *EURASIP Journal on Applied Signal Processing*, vol. 2005, pp. 1400–1409, 2005.

[9] J. Kim, M. Nasir, R. Gupta, M. Van Segbroeck, D. Bone, M. Black, Z.I. Skoridlis, Z. Yang, and S. Narayanan, "Automatic estimation of Parkinson's disease severity from diverse speech tasks," in *Proc. Interspeech*, 2015.

[10] . Kim, N. Kumar, A. Tsiartas, M. Li, and S.S. Narayanan, "Automatic intelligibility classification of sentence-level pathological speech," *Computer speech & language*, vol. 29, no. 1, pp. 132–144, 2015.

[11] G. An, D.G. Brizan, M. Ma, M. Morales, A.R. Syed, and A. Rosenberg, "Automatic recognition of unified Parkinson's disease rating from speech with acoustic, i-vector and phonotactic features," in *Proc. Interspeech*, 2015.

[12] C. Middag, T. Bocklet, J.P. Martens, and E. Nöth, "Combining phonological and acoustic asr-free features for pathological speech intelligibility assessment.," in *Proc. Interspeech*, 2011, pp. 3005–3008.

[13] A. Zlotnik, J.M. Montero, R. San-Segundo, and A. Gallardo-Antoln, "Random forest-based prediction of Parkinson's disease progression using acoustic, ASR and intelligibility features," in *Proc. Interspeech*, 2015.

[14] A. Maier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster, and E. Nöth, "PEAKS - A system for the automatic evaluation of voice and speech disorders," *Speech Communication*, vol. 51, no. 5, pp. 425–437, 2009.

[15] H.V. Sharma, M. Hasegawa-Johnson, J. Gunderson, and A. Perlman, "Universal access: Preliminary experiments in dysarthric speech recognition," in *Proc. Interspeech*, 2009, p. 4.

[16] C. Middag, J.P. Martens, G. Van Nuffelen, and M. De Bodt, "Automated intelligibility assessment of pathological speech using phonological features," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 3, 2009.

[17] D. Duez, "Consonant and vowel duration in parkinsonian french speech," in *Proceedings of Speech Prosody*, 2006, pp. 101–105.

[18] Johanna K Lake, Karin R Humphreys, and Shannon Cardy, "Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders," *Psychonomic bulletin & review*, vol. 18, no. 1, pp. 135–140, 2011.

[19] P.A. Heeman, R. Lunsford, E. Selfridge, L. Black, and J. Van Santen, "Autism and interactional aspects of dialogue," in *Proc. Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2010, pp. 249–252.

[20] R. Gupta, P.G. Georgiou, D. Atkins, and S.S. Narayanan, "Predicting client?s inclination towards target behavior change in motivational interviewing and investigating the role of laughter," in *Proc. InterSpeech*, 2014.

[21] D. Bone, T. Chaspari, K. Audhkhasi, J. Gibson, A. Tsiartas, M. Van Segbroeck, M. Li, S. Lee, and ShS.rikanth Narayanan, "Classifying language-related developmental disorders from speech cues: the promise and the potential confounds.," in *Proc. Interspeech*, 2013, pp. 182–186.