

## Accepted Manuscript

On the Issues of Intra-Speaker Variability and Realism in Speech,  
Speaker, and Language Recognition Tasks

John H.L. Hansen, Hynek Bořil

PII: S0167-6393(17)30384-9  
DOI: [10.1016/j.specom.2018.05.004](https://doi.org/10.1016/j.specom.2018.05.004)  
Reference: SPECOM 2566



To appear in: *Speech Communication*

Received date: 13 October 2017  
Revised date: 17 May 2018  
Accepted date: 24 May 2018

Please cite this article as: John H.L. Hansen, Hynek Bořil, On the Issues of Intra-Speaker Variability and Realism in Speech, Speaker, and Language Recognition Tasks, *Speech Communication* (2018), doi: [10.1016/j.specom.2018.05.004](https://doi.org/10.1016/j.specom.2018.05.004)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# On the Issues of Intra-Speaker Variability and Realism in Speech, Speaker, and Language Recognition Tasks<sup>☆</sup>

John H. L. Hansen<sup>a,\*</sup>, Hynek Bořil<sup>a,b</sup>

<sup>a</sup>*Center for Robust Speech Systems (CRSS), Erik Jonsson School of Engineering, University of Texas at Dallas, Richardson, Texas, U.S.A.*

<sup>b</sup>*Pioneer Speech Signal Processing Laboratory (PSSPL), Electrical and Computer Engineering Department, University of Wisconsin-Platteville, U.S.A.*

---

## Abstract

Recent years have witnessed notable advancements in the areas of speech, speaker and language/dialect recognition. However, many of the emerging scientific principles appear to be drifting to the sidelines with the assumption that access to larger amounts of data is all that is required to address a growing range of issues relating to new scenarios. This study surveys several challenging domains in formulating effective solutions in realistic speech data, and in particular the notion of using naturalistic data to better reflect the potential effectiveness of new algorithms. Our main focus is on intra-speaker mismatch and speech variability issues due to (i) differences in noisy speech with and without Lombard effect and a communication factor, (ii) realistic field data in noisy and increased cognitive load conditions, (iii) speech variability introduced by whispered speech, and (iv) dialect identification using found data. Finally, we study speaker–environment and speaker–speaker interactions in a newly established, fully naturalistic Prof-Life-Log corpus. The specific outcomes from this study include an analysis of the strengths and weaknesses of simulated vs. actual speech data collection for

---

<sup>☆</sup>This project was funded in part by AFRL under contract FA8750-15-1-0205 and partially by the University of Texas at Dallas from the Distinguished University Chair in Telecommunications Engineering held by J. H. L. Hansen.

A preliminary short version of this article was published in the special session “Realism in Robust Speech Processing” in ISCA INTERSPEECH 2016 [1].

\*Corresponding author

*Email address:* [John.Hansen@utdallas.edu](mailto:John.Hansen@utdallas.edu) (John H. L. Hansen)

*URL:* <http://www.utdallas.edu/~john.hansen> (John H. L. Hansen)

Download English Version:

<https://daneshyari.com/en/article/6960485>

Download Persian Version:

<https://daneshyari.com/article/6960485>

[Daneshyari.com](https://daneshyari.com)