

Title	ロンバード効果に着想を得た雑音中での音声了解度および自然性の向上
Author(s)	NGO, THUANVAN
Citation	
Issue Date	2020-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/16995
Rights	
Description	Supervisor:赤木 正人, 先端科学技術研究科, 博士

Abstract of Doctoral Dissertation

**Improvement of intelligibility and naturalness of speech in
noise inspired by Lombard effect**

For doctor degree of information science

Thuanvan Ngo
1720010

Supervisor: Professor Masato Akagi
Akagi Laboratory

Graduate School of Advanced Science and Technology

Japan Advanced Institute of Science and Technology

Information Science

September 2020

Research content and purpose

In public announcements in train stations or airports, the presence of noise often smears presented speech, thus makes it hard for listeners to understand it. By reducing noise throughout the presentation, speech is still intelligible and natural to listeners. However, this seems impractical due to complex architectures of these places and costly installed devices. Besides, it is practical and less expensive to enhance speech itself before presenting to complement degradation in intelligibility and naturalness by the smears.

Lombard speech is intelligible speech produced by humans in noise due to the Lombard effect. Investigation of Lombard speech could reveal essential features to increase speech intelligibility and naturalness. Therefore, the purpose of this research was to improve the intelligibility and naturalness of speech in noise using conversion rules inspired by the Lombard effect. It came up with two sub-goals: (I) obtaining feature understanding and control which contributes to the intelligibility of Lombard speech under **noise-level-varying and various noise** and (II) identifying and applying the **effective feature control methods** for exceeding the intelligibility and naturalness of Lombard speech. From the previous research and the properties of Lombard speech that varies with noise levels and noise types, three problems arose to cover the search space (features, noise levels, feature variations, SNRs, and spectral-varied noise) for finding features and applying them.

(1) Contribution of acoustic features of Lombard speech has no consideration of their articulatory features at one noise level: For this problem, the modification of acoustic features of Lombard speech was often done without any considering the articulatory changes or challenging to be obtained in acoustical levels. The multiple contributions of the features to the intelligibility and naturalness were unclear.

(2) Control and contribution of acoustic features of Lombard speech in multiple noise levels of backgrounds: In this problem, acoustic features contributing to Lombard speech in various noise levels were difficult to be modeled and controlled by the conventional methods.

(3) Unclear effective features to the intelligibility and naturalness of speech varying noise levels and various types of noise: This problem followed the second problem. Acoustic features, when varying that contribute to the intelligibility of speech in noise, remained unclear. Recent studies differently reported effective features for the intelligibility and naturalness of speech in noise. Thus, the precise set of effective features was unidentified.

Thus, the investigation has three steps: (1) Mimicking Lombard speech by controlling articulatory and acoustic features, (2) Effective features for the intelligibility and naturalness of speech in noise, (3) Application to improve the intelligibility of speech under noisy reverberant conditions.

Following these steps, this study obtained the articulatory and acoustical controls in mimicking Lombard speech. The contributive features, including spectral tilts, f_o , and formants, were also explored in the first step. In the second step, the effective features for intelligibility and naturalness in all kinds of noise were identified. In the final step, the effective features were successfully applied to increase intelligibility and naturalness of speech under noisy reverberant conditions with a pair of effective time-frequency features.

Consequently, the originality of this study was firstly the leading investigation of the contribution of articulatory features to the intelligibility of speech in noise. Secondly, it was mimicking Lombard speech under various noise levels. Finally, this study presented a brutal-force method for extracting the effective acoustic features to vary to increase the intelligibility of speech in noise. Besides, the present study obtained novelty in two aspects. Firstly, it was the concept of applying rule-based methods and the Lombard effect model for the rule generation model to mimic Lombard speech concerning multiple noise levels. Secondly, it was the concept based on the modulation spectrum and modulation transfer function concepts in relationship with listening tests to identify the effective features to increase speech intelligibility and naturalness in noise.

As a scientific orientation, this research can enlighten the fields of speech enhancements, objective intelligibility measurements, voice conversion, and synthesis. Especially, it provides essential, necessary information for the areas of speech enhancement engineering.

Keywords: Lombard speech, mimicking, articulatory features, acoustic features, effective features, speech intelligibility and naturalness in noise.

Research accomplishment (publications)

Journal (Peer reviewed)

- [1] **Thuanvan Ngo**, Rieko Kubo, Daisuke Morikawa and Masato Akagi, “Acoustical Analyses of Tendencies of Intelligibility in Lombard Speech with Different Background Noise Levels,” *Journal of Signal Processing*, vol. 21, no. 4, pp. 171–174, 2017.
- [2] **Thuanvan Ngo**, Masato Akagi, and Peter Birkholz, “Effect of articulatory and acoustic features on the intelligibility of speech in noise: An articulatory synthesis study,” *Speech Communication*, vol. 117, pp. 13–20, 2020.
- [3] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Mimicking Lombard effect: An analysis and reconstruction,” *IEICE Trans. Inf. & Syst.*, vol. E103.D, no. 5, pp. 1108–1117, 2020.

International conference (Peer reviewed)

- [4] **Thuanvan Ngo**, Rieko Kubo, Daisuke Morikawa and Masato Akagi, “Acoustical analyses of Lombard speech by different background noise levels for tendencies of intelligibility,” *NCSP’17*, 2017, pp. 309–312.
- [5] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Evaluation of the Lombard effect model on synthesizing Lombard speech in varying noise level environments with limited data,” *APSIPA ASC*, 2019, pp. 133–137.

Domestic conference (Non peer reviewed)

- [6] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Intelligibility improving and naturalness preserving for evacuation speech in noisy environments,” *IEICE Technical Report, Engineering Acoustics*, 2019.
- [7] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Improved quality and intelligibility of mimicking Lombard speech by source-filter and coarticulation model-based synthesis,” *ASJ Spring Meeting*, 2019.
- [8] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Speaker-independent control model for mimicking Lombard speech uttered in background noises with various levels,” *ASJ Spring Meeting*, September, 2018, pp. 1371-1374.
- [9] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Acoustical control method for increasing intelligibility based on Lombard speech uttered in background noises with various levels,” *ASJ Fall Meeting*, 2018, pp. 313-316.
- [10] **Thuanvan Ngo**, Rieko Kubo, and Masato Akagi, “Acoustical rules for mimicking Lombard speech produced in a various noise level background,” *IEICE Technical Report, Engineering Acoustics*, vol. 117, no. 170, 2017.