Speech Intelligibility Improvement in Car Noise Environment by Voice Transformation

Karan Nathwani, Gaël Richard, Bertrand David, Pierre Prablanc, Vincent Roussarie

PII: S0167-6393(16)30194-7
DOI: 10.1016/j.specom.2017.04.007
Reference: SPECOM 2457

To appear in: *Speech Communication*

Received date: 27 July 2016
Revised date: 3 February 2017
Accepted date: 25 April 2017

Please cite this article as: Karan Nathwani, Gaël Richard, Bertrand David, Pierre Prablanc, Vincent Roussarie, Speech Intelligibility Improvement in Car Noise Environment by Voice Transformation, *Speech Communication* (2017), doi: 10.1016/j.specom.2017.04.007

Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 00 (2017) 1–16

Speech Communi-cation

Speech Intelligibility Improvement in Car Noise Environment by Voice Transformation

Karan Nathwani[1], Gaël Richard[1], Bertrand David[1], Pierre Prablanc[2] and Vincent Roussarie[2]

1: LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France

2: PSA Peugeot Citroën, Chemin de Gisy, 78943 Vélizy-Villacoublay

Abstract

The typical application targeted by this work is the intelligibility improvement of speech messages when rendered in car noise environment (radio, message alerts,...). The main idea of this work is to transform the original speech to "Lombard" speech or more precisely to simulate some of the strategies followed by humans to render their speech clearer when they are surrounded by noise. Three main effects are considered in this work, namely non uniform-time scale modification, formant shifting and a combination of these modifications along with energy redistribution between speech regions. All effects are studied with specific transformations for voiced and unvoiced segments. The proposed modifications are then evaluated by means of subjective and objective tests. The results of these tests conducted with normal hearing and impaired listeners demonstrate the potential of the selected transformations for voice intelligibility improvement.

Keywords: Speech Intelligibility, Car Noise Environment, Hearing in Noise Test, Voice Transformation, Lombard Speech.

1. Introduction

Speech Intelligibility usually refers to a measure of the effectiveness of understanding speech. It is used to evaluate telecommunications systems performances, to characterize some acoustical properties of conference rooms or to evaluate the level of understanding of patients for medical purposes.

It is well known that intelligibility is affected by the presence of background noise which can mask crucial portions of the speech content. It is also widely acknowledged that humans succeed to enhance the audibility of their voice by means of a number of non-linear effects which are often gathered under the term of Lombard effect [1, 2, 3, 4, 5]. In fact, humans do modify their speech in the presence of noise in such a way to enhance the acoustic contrasts between their speech and the background noise. It is for example shown in [6] that besides a straightforward strategy of speaking louder, the speakers did alter other vocal characteristics such as for example the center frequency of the first formant or the modulation of the fundamental frequency on voiced segments.

For many applications including speech recognition or telephony in adverse conditions it is necessary to reduce the impact of the surrounding noise and to mitigate channel effects. This explains the vast literature on speech enhancement and dereverberation [7, 8]. However, improving speech quality does not necessarily improve speech intelligibility [9, 10] and factors explaining why speech quality is not directly linked to speech intelligibility are studied but not yet well understood [11]. This may be explained by the fact that speech enhancement algorithms maximize a cost function which may not be well correlated with speech intelligibility.

In this work, we focus on a different application, namely the improvement of the intelligibility of a spoken message (possibly originally uttered in a quiet environment) when rendered in a noisy environment. A typical application concerns the intelligibility improvement of speech messages in the car environment (radio, alert messages, telephony,...).