Technical University of Denmark

DTU

Listening talkers produce great spectral tilt contrasts

Christiansen, Thomas Ulrich; Heegård, Jan; Henrichsen, Peter Juel

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Listening talkers produce great spectral tilt contrasts

Thomas Ulrich Christiansen[1], Jan Heegård[2], Peter Juel Henrichsen[2]

[1]Centre for Applied Hearing Research, Department of Electrical Engineering,
Technical University of Denmark, Lyngby, Denmark
[2]Department of International Language Studies and Computational Linguistics, Copenhagen Business
School, Copenhagen, Denmark

tuc@elektro.dtu.dk, jhp.isv@cbs.dk, pjh.isv@cbs.dk

Abstract

It is well known that the envelope of the long-term average speech spectrum flattens with vocal effort. A recent study [1] showed that content words had a flatter spectral envelope than content words at the same overall level for a specific Danish speech material.

The present paper investigates whether this effect is present in a larger and more diverse speech material, and if the effect is greater when the talker is listening (participating in a dialogue) as compared to monologue.

The monologue speech material consisted of recordings from 18 native talkers of Danish describing a network of colored geometrical shapes taken from DanPASS [2]. The spectral tilt was gauged by calculating the band-level difference in dB between two frequency bands with pass-bands 150 to 803 Hz and 803 to 1358 Hz respectively in 5 ms intervals.

This was done separately for intervals containing content words and function words and grouped by talker. The spectral tilt difference was then calculated as the average band-level difference for function words minus the average band-level difference for content words. This calculation was grouped per talker. For the monologues these differences ranged between 5 and 8 dB for the 18 talkers.

Content words were defined as nouns, active verbs, adjectives and adverbs. Function words were defined as articles, pronouns, conjunctions and auxiliary verbs. Words not belonging to any of these categories were not used.

The dialogue speech material was also from DanPASS and consisted of recordings from 13 of the same talkers as the monologues. In the dialogue speech material talkers where asked to describe a map with certain discrepancies and negotiate their way through the map.

Spectral tilt differences between content- and functions words were calculated in the same way as for the monologues. The results show that the spectral tilt differences are slightly higher for dialogues than monologues. A two-way anova (grouped by talker and word type) showed that these differences are significant.

We conclude that Danish talkers mark high information density in spontaneous speech (=content words) by means of flat spectral envelope, not just for monologues, but also for dialogues. Moreover, when engaged in dialogue, talkers enhance this spectral flattening.

In our view it is remarkable that conclusions with statistical validity can be reached based on the over-simplified definition of spectral tilt employed in this paper. We speculate that optimizing both the definition of spectral tilt and the word categories comprising content- and function words, may allow us to observe even greater effects than reported here.

The eventual goal of this line of research is to devise a simple, tractable method for distinguishing high information content from low information content in speech, based on the ubiquitous assumption that content words carry more information than function words. Such a method could potentially be applied in hearing aids, cochlear implants and automatic speech recognition.

Index Terms: spectral tilt, spectral envelope, speech production, speech perception, content words, function words.

References

[1] Henrichsen, P.J. and Christiansen, T.U., "Information based speech transduction", Proceedings of 3rd International Symposium on Auditory and Audiological Research, August 2011, Nyborg, Denmark, in press, 2011

[2] Grønnum, N., "A Danish phonetically annotated spontaneous speech corpus (DanPASS)", Speech Communication 51, 594-603, 2009