Pure

Syddansk Universitet

Virtual reality simulations as a new tool for practicing presentations and refining public-speaking skills

Niebuhr, Oliver; Michalsky, Jan

Virtual reality simulations as a new tool for practicing presentations and refining public-speaking skills

Oliver Niebuhr[1], *Jan Michalsky*[2]

[1]SDU Electrical Engineering, Mads Clausen Institute, University of Southern Denmark
[2]Institute of German Studies, University of Oldenburg, Germany
olni@sdu.dk, j.michalsky@uol.de

Abstract

Presentations are typically practiced alone while talking to oneself in a silent room. It is not only questionable whether such a rehearsal setting is a proper preparation for a real public-speaking situation. Giving the same talk repeatedly to oneself also bears the risk that speaking "erodes" from the communicative act of conveying a message into a mere mechanical exercise that is neither content- nor audience-oriented. Against this background, it is tested from a digital-humanities perspective whether a VR public-speaking simulation, in which a speaker can rehearse his/her talk in a virtual conference room and in front of a virtual audience, is a suitable and preferable alternative to practicing a presentation on one's own. Prosodic measures of speaking style are analyzed and compared between two groups of 12 speakers, a control group and a VR test group, each of which performed several rounds of practicing. Results suggest that test-group speakers take the VR environment seriously and show, unlike control group speakers, an audience-oriented, more charismatic speaking style, with reduced signs of prosodic erosion due to repeated rehearsal. These findings are discussed in the light of digital-humanities applications of VR technology.

Index Terms: prosody, public speaking, virtual reality.

1. Introduction

1.1. Background and questions

Speaking in front of an audience is an integral part of every-day professional life. It already starts at school with giving presentations for the other class mates, and in some professional careers it culminates in holding keynote speeches for hundreds or thousands of customers, investors, share and/or stake holders. The resulting need to practice presentations and refine public-speaking skills is reflected in the promotion of debate clubs at schools and in the many rhetoric-training institutions, companies, and textbooks that help politicians, managers, entrepreneurs, teachers, and sales agents become more memorable and charismatic speakers [1]. Steve Jobs was known for refining and rehearsing his presentations "endlessly and fastidiously" [38]. The results were seemingly effortless and exceptionally charismatic keynote speeches. Rhetorical training is literally worth millions in terms of its annual financial turnover and with respect to clicks on YouTube [2,3].

For some years now, there is a technology available that simulates presentation situations in a Virtual Reality environment [4]. This technology, which is a prime example of applied digital humanities, seems to have a high potential for practicing presentations and refining public speaking skills, but this potential has remained widely untested so far.

Against this background, the present paper aims to shed some initial light on two questions. The first question (Q1) is a proof-of-concept question. Obviously, VR simulations of presentation situations can only be an effective tool for practicing talks and refining public-speaking skills if speakers take the virtual environment seriously and at face value. This would mean to treat the virtual audience in the simulated presentation situation as if it was a real audience and this, in turn, would mean that the prosodic differences between giving a talk for oneself and giving a talk in front of a virtual audience resemble the prosodic differences between monologue and dialogue speaking styles. Moreover, when talking to oneself, the speaker-listener distance is virtually zero, whereas talking in a VR presentation scenario means addressing listeners some of which are a few (virtual) meters away from the speaker. Thus, if speakers take this virtual distance for real, then they will take measures for their voices to fill the entire room and reach, on an equal level, all listeners whom they are talking to.

In summary, we would expect in connection with (Q1) that speakers in VR presentation scenarios spontaneously and consistently adopt an addressee-oriented (rather than a self-oriented) way of speaking, which includes adjusting the vocal effort such that it bridges a greater speaker-listener distance.

The second question (Q2) builds upon the first and relates to the fact that practicing presentations always means repeated rehearsal. However, without somebody to talk to, the supposed-to-be communicative action of giving a talk can quickly erode into a mere mechanical exercise in which the speaker gets bored and loses the intention of actually getting a message across. Such an exercise may still help the speaker memorize the talk's individual words and sentences, but beyond that it does not leave him/her with a good representation of how an attention-attracting, charismatic performance should sound like. The question (Q2) is whether the VR presentation environment is able to reduce or even block this rehearsal-related erosion process. Such a positive effect of VR technology on presentation rehearsal is expectable in particular if the VR audience is taken seriously and at face value by speakers.

1.2. Prosodic parameters related to (Q1) and (Q2)

With respect to the proof-of-concept question (Q1), studies that specifically examined the prosodic differences between monologues and dialogues are scarce [5,6]. But, since read speech has been mainly investigated in monologues and spontaneous speech in dialogues, we can assume that some of the acoustic differences characterizing read speech and spontaneous speech are connected to the presence of an interlocutor.

Although read speech and spontaneous speech both constitute a spectrum of styles depending on medium, preparation, and the amount of interaction [7,8,9,10,11], there

are recurring acoustic features associated with the differences between spontaneous and read speech. Compared to read-speech monologues, spontaneous speech dialogues show, for example, a lower speaking rate [5,12,13,14], a higher F0 mean [9] and a larger F0 variability [6,9,12,15,16].

Concerning the vocal-effort part of (Q1), research on an increased speaker-listener distance found that speakers increase both their intensity and their F0 levels under these circumstances [17,18]. Thus, the F0 level is involved in both a listener-oriented spontaneous speaking style and an increased vocal effort associated with the intention of a speaker to reach remote listeners with his/her voice and message.

With respect to the erosion-by-rehearsal question (Q2), [19] showed on the basis of the traditional sentence-list elicitation paradigm of laboratory phonetics that the monotony of this task in combination with the repetition of sentences during the experiment (even if not in direct succession) makes speakers increase their speaking rate, lower their F0 level, and narrow their F0 range. In other words, the speakers' intonation became more monotonous and their articulation faster and hence more sloppy. Effective non-erosive rehearsal of speech, on the other hand, minimizes these changes and makes speakers additionally talk more fluently which includes, for example, longer phrases and fewer interrupting pauses or disfluencies [20].

1.3. Prosodic expectations for (Q1) and (Q2)

Regarding the motivation of our study, one may wonder whether it is at all realistic to expect that adults change their mental and physical behavior just because they take on a pair of VR glasses. The answer to this implicit question is yes, it is realistic. For example, findings of previous studies conducted in the realm of psychology found that talking in front of a virtually simulated audience was able to first trigger and then successively reduce the speakers' anxiety of public speaking. This was true, although all participants were fully aware of the virtual nature of the audience [21]. Furthermore, talking in front of a positive, neutral, or aggressive virtual audience significantly affected the speakers' anxiety level and its reduction in a positive or negative way [22], see also [34,35,36].

These are clear indications that speakers can automatically and against their better knowledge treat virtual audiences in simulated presentation situations like real audiences and at face value. Based on these indications from anxiety research, we expect the following prosodic effect to occur in connection with our (Q1) and (Q2).

Compared to a control group of speakers who practice their presentation on their own in a silent room, a test group of speakers who can practice their talk within a VR presentation scenario (see top right corner in Fig.2) will...

- (Q1): show a lower speaking rate, a higher F0 level and range, and -- due to an increase in vocal effort that meets the greater speaker-listener distance -- also a higher intensity level.

- (Q2): show, over the repetitions of their talk, no or reduced speech-erosion changes, whereas such changes cause the control group to successively lower their F0 level, narrow their F0 range, increase their speaking rate, and -- due to the fading relevance of the speaker-listener distance -- also lower their intensity level. The test group's speech will instead become more fluent, for example, in the form of longer prosodic phrases and fewer pauses or disfluencies.

2. Method

2.1. Speakers

Twenty-four male post-graduate students of the Innovation & Business Msc program at the University of Southern Denmark Sonderborg (SDU) participated in the study. The average age was 23.5 years. They were all L2 speakers of English at the top proficiency levels C1 or C2. Furthermore, they all had some presentation experience and stated to feel comfortable in public speaking scenarios. Thus, anxiety of public speaking was excluded as a confounding factor in our experiment.

2.2. Procedure

The 24 participants were randomly assigned to two equally large groups, a control group and a VR test group. The 12 participants in each group were instructed to prepare an "elevator pitch" based on the same set of three given powerpoint slides. The slides were about the development of a new bio-tech filter that would sustainably improve air quality. An "elevator pitch" is a concise 2-3 minute presentation of a business idea [37]. Elevator pitches are given with the aim to make people invest in the speaker's company or to give the speaker's company access to facilities of business networks, municipalities, etc.

All participants had learned to develop and hold such persuasive elevator pitches as part of their study program; and, on this basis, they were given about five minutes time (in a separate silent lecture room) to familiarize themselves with the slide deck that was handed out to them.

After the familiarization phase, the members of both groups were individually instructed to practice the elevator pitch four times in a row, with the aim to be able to present it afterwards in front of an audience of international company representatives who visited the SDU at that day and were interested in the university's research and commercialization activities. The participants were not informed at that stage (but later on in the de-briefing) that this aim was just a pretext to put the speakers in a realistic preparation context and make them rehearse the elevator pitch conscientiously. Participants of the control group were asked to do this rehearsal aloud, while imagining to address real listeners and seek their support for the pitched idea. The slides were uploaded to an interactive TV screen in front of them. In addition, participants were asked to take a break of about 10 minutes in between two rehearsals and to wear a headset microphone in order to make the rehearsal more realistic. They were not informed at this stage that their rehearsals were recorded through the headset, but during the de-briefing all speakers gave their written consent for the recorded speech data to be stored and analyzed. The control group's rehearsals were recorded digitally at 44.1 kHz/16-bit in a sound-attenuated lecture room at SDU.

For the test group, the procedure was exactly the same, except that the test-group participants were asked to do their rehearsals within a virtual environment. Besides that the test-group participants also had to address the (virtual) listeners while speaking, and they were, like the control-group participants, asked to present persuasively as if they would really seek the listeners' support for their pitched idea. All VR group participants had gained some experience with VR applications during lectures at SDU. So, we can assume that the obtained data is free from any technology familiarization artifacts.

The virtual simulation projected the participants within a large lecture hall in front of a simulated audience of about 20

listeners, see Figure 1, top panel. As can also be seen in the bottom panel of Figure 1, the elevator-pitch slide deck was directly uploaded into the virtual environment and projected by a virtual beamer onto a virtual screen so that the participants were able to see and click through the slides like in a real-world presentation. Furthermore, the virtual audience gave neutral responses such as breathing and coughing to enhance the speaker's immersion experience [22]. The VR software we used is called "Presentation Simulator" [23]. The VR simulation was running on a Windows 10 desktop PC. The VR interface was a HTC Vive headset including controllers.
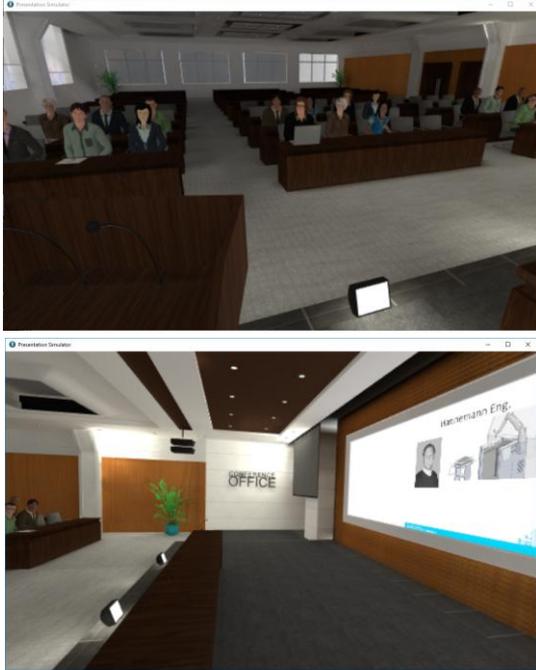

Figure 1: *Simulated audience and conference hall in which the VR test-group participants rehearsed their elevator pitch.*

2.3. Acoustic analysis

Relevant features were automatically extracted from the recordings, using the existing PRAAT [24] scripts described in the works of [25] and [26]. The extracted features included three counts and four acoustic measures. The counts were:

- The total number of syllables per elevator pitch, estimated on the basis of intensity maxima.
- The total number of pauses, i.e. silences ≥ 300 ms.
- The number of syllables per inter-pausal unit (IPU).

The measured parameters were:

- The average speaking rate (syl/s) per IPU,
- the average F0 level (in Hz) per IPU,
- The F0 range (in semitones, st) per IPU.
- The intensity level (in dB) per IPU.

The sample sizes were n = 12 for the first two counts and varied between n=600-900 for the measurements taken per IPU (all 12 elevator pitches contained 50-70 IPUs).

2.4. Statistical analysis

Note that (Q1) is tested by comparing the control-group data with the test-group data, whereas (Q2) is tested by comparing the data among the four rehearsals in each group condition. Thus, the (Q1) comparisons are between-subjects comparisons, whereas the (Q2) comparisons are within-subjects comparisons. Accordingly, we used mixed-design GLMs (General Linear Models) for our statistical analysis. Fixed factors were VR and REHEARSAL. Dependent variables were the counts or mean values per subject (n=12). Subject (SPEAKER) was included as a covariate.

3. Results

A descriptive results summary is provided in Figures 2(a)-(g). Regarding the combined influences of VR and REHEARSAL on the test group's and control group's presentation behavior, the overall results picture shows that the various acoustic features fall into three different patterns. Repeatedly rehearsing the elevator pitch...

- (i) had the same effect on both groups, but this effect was more strongly pronounced in the control group than in the test group. This unidirectional VR*REHEARSAL interaction pattern occurred for the total number of syllables ($F[3,63]=8.9$, $p<0.001$, $\eta_p^2=.30$) and the speaking rate ($F[3,63]=14.2$, $p<0.001$, $\eta_p^2=.40$).

- (ii) had no effect on the VR test group, but a strong effect on the control group. This significant interaction pattern applied to the F0 range ($F[3,63]= 16.9$, $p<0.001$, $\eta_p^2=.45$).

- (iii) had similarly strong but opposite effects on both groups. This interaction pattern was indeed the most frequent one. The corresponding VR*REHEARSAL significances concerned the total number of pauses ($F[3,63]=26.7$, $p<0.001$, $\eta_p^2=.56$), the number of syllables per IPU ($F[3,63]=9.6$, $p<0.001$, $\eta_p^2=.32$), the F0 level ($F[3,63]=15.2$, $p<0.001$, $\eta_p^2=.42$) and the intensity level ($F[3,63]=6.6$, $p=0.001$, $\eta_p^2=.24$).

In addition to these three interaction patterns, we found several significant main effects of VR. These main effects show that VR test group participants produced more speech than control group participants ($F[1,21]=219.4$, $p<0.001$, $\eta_p^2=.91$), that the control group's speech, in turn, was more often interspersed with pauses ($F[1,21]=85.4$, $p<0.001$, $\eta_p^2=.80$), and that, in consequence, the test group's IPUs were longer than those of the control group, i.e. less frequently interrupted by pauses and other disfluency phenomena ($F[1,21]=265.6$, $p<0.001$, $\eta_p^2=.93$). Furthermore, VR test group participants gave the elevator pitch on higher F0 and intensity levels ($F[1,21]=72.5$, $p<0.001$, $\eta_p^2=.78$; $F[1,21]=13.4$, $p=0.001$, $\eta_p^2=.39$), with a larger F0 range ($F[1,21]=27.1$, $p<0.001$, $\eta_p^2=.56$), and at a slower speaking rate ($F[1,21]=5.3$, $p=0.03$, $\eta_p^2=.12$) than control group participants.

Main effects of REHEARSAL were restricted to the unidirectional pattern-(i) interactions. The total number of syllables decreased (i.e. elevator pitches got shorter) across the four rehearsals ($F[3,63]=9.3$, $p<0.001$, $\eta_p^2=.31$), whereas the speaking rate increased successively ($F[3,63]=4.2$, $p<0.01$, $\eta_p^2=.17$).

Finally, the covariate SPEAKER had a significant main effect on none of the dependent variables nor did it interact with the two factors VR and REHEARSAL. We interpret this as showing that we indeed succeeded with forming two homogeneous speaker groups for our VR and control conditions (e.g., by controlling speaker age, educational background as well as public-speaking and VR experience).
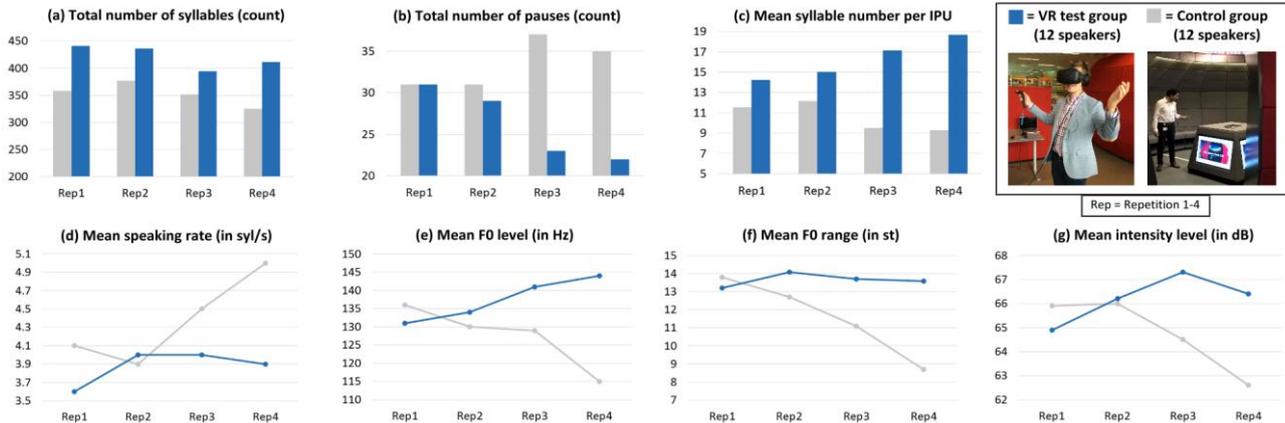
Figure 2: *Results summary of the acoustic analysis.*

4. Discussion

In the realm of phonetics, digital humanities refer to the use of digital technologies and devices for storing and analyzing but also supporting and changing speech patterns. In the latter context, we tested in our study the suitability of VR public-speaking simulations for practicing presentations like elevator pitches. Two questions were addressed. (Q1) was a proof-of concept question: Do speakers take the VR scenario, including the simulated audience, for real and at face value? (Q2) was an erosion-by-rehearsal question (Q2): Can the speech deterioration phenomena that were found to occur when speakers repeat the same message over and over again be reduced or blocked by practicing presentations in the VR scenario?

Comparing speakers of a control group to speakers of a VR test group suggest that the answers to both questions are positive. In accord with our prosodic expectations, we found evidence that the speakers of the VR test group gave their elevator pitches in a more listener-oriented, dialogue speaking style than the speakers of the control group. Compared to the latter group, the presentations of the VR test group were characterized by a higher F0 level, a larger F0 range, and a slower speaking rate [6,9,12,13,14,15,16]. Moreover, matching the relatively large speaker-listener distance in the VR situation, the VR test group speakers gave their elevator pitches at a higher intensity level [17,18]. Presenting in a VR environment also unexpectedly animated the speakers to talk longer. With respect to the speech deterioration, our control group speakers indeed showed a successive lowering and narrowing of their speech melody across the repeated rehearsals of their presentation [19]. In addition, their speech became faster and less fluent. Compared to that, the VR test group speakers did not show these signs of speech erosion at all, or to a lesser degree, or even changed their speech into the opposite direction, thus approaching parameter settings that are known to be associated with increased speaker charisma [27,28,29,30,31,32].

In summary, the virtual audience in a VR simulation causes a response in speaking style comparable to that of a physically present audience. Furthermore, the participants of the control group showed an overall decrease in presentation performance, whereas the VR group participants even improved their presentation performance to some degree in the course of the four repetitions. Thus, our results suggest that speaking in a VR scenario and in front of a VR audience is can be a useful tool in various digital-humanities applications. These applications range from more practical to more scientific ones. On the practical side, there is the presentation and

public-speaker training that served as starting point of the present study. We currently conduct three follow-up studies in this line of research. First, we put speakers in different presentation scenarios (e.g., smaller and larger meeting rooms with smaller and larger audiences) and also add a real-life presentation situation at the end of the rehearsal process in order to get a better understanding of how VR variables influence the prosodic characteristics of a speaker's presentation and test whether these influences are lasting in the sense that they are carried over into a subsequent real-life presentation situation. Second, we copy real lecture and conference rooms at SDU (as well as corresponding sample audiences) into the VR simulation. This enables us to compare speech and presentation behavior in exactly the same settings inside and outside VR and, thus, to examine the proof-of-concept question (Q1) at a new level of detail and ecological validity. Third, we pick up on the indications of the present results that VR technology is not just an effective means of listener-oriented, non-erosive presentation training, but also beneficial in that it could stimulate presenters to adopt a more charismatic speaking style [29,31]. Further acoustic analyses and complementary perception experiments are run to pursue this VR potential further.

On the scientific side, the general context-sensitivity of speakers to VR environments suggests that VR technology will prove very useful in laboratory phonetics, for example, when it comes to bridging the controversially discussed gap between making high-quality speech recordings on the one hand and eliciting everyday conversational or expressive speech under "field-like" conditions on the other [33]. One big innovation in this context would be the development of interactive addressees or audiences within VR environments. For example, we currently develop a VR interface that evaluates a speaker's presentation performance based on acoustic-prosodic parameter analyses and then adjusts the visible audience behavior accordingly on a attentive-distracted scale. Such innovations are, in the end, not just something that phonetic researchers can benefit from, they also represent a new important field of research that phoneticians can contribute to.

5. Acknowledgments

6. References

[1] O. Cabane Fox, *The charisma myth: How anyone can master the art and science of personal magnetism.* New York: Penguin, 2012.

[2] P. Sterman, "A toast to toastmasters! - On the organization's 90th anniversary, the future looks brighter than ever," In *Toastmaster - The Magazin for Communicators and Leaders Oct.2014*, pp. 16-21, 2014.

[3] J. Treasure, "How to speak so that people want to listen," retrieved 11 November 2017 from *https://www.youtube.com/watch?v=eIho2S0ZahI&t=8s*, 2014.

[4] D. Kariuki, "Five VR apps that help you be a better public speaker," Hypergridbusiness, retrieved 11 November 2017 from *http://www.hypergridbusiness.com/2016/08/5-vr-apps-that-help-you-be-a-better-public-speaker/* , 2016.

[5] O. Niebuhr, J. Bergherr, S. Hutz, C. Lill and J. Neuschulz, "Intonationsfragen hinterfragt - die Vielschichtigkeit der prosodischen Unterschiede zwischen Aussage- und Fragesätzen mit deklarativer Syntax," *Zeitschrift für Dialektologie und Linguistik*, vol. 77, pp. 304–346, 2010.

[6] K. Görs and O. Niebuhr, "Hocus Focus - What the Eliciation Method Tells Us About Types and Exponents of Contrastive Focus," *Proceedings of Speech Prosody 6, Shanghai, China*, pp. 262–265, 2012.

[7] W. Labov, "The isolation of contextual styles," In: W. Labov (eds.). *Sociolinguistic patterns*, Oxford. Blackwell, 1972, pp. 70–109.

[8] B. Granström, "The use of speech synthesis in exploring differentspeaking styles," *ESC ETRW on the Phonetics and Phonology of Speaking Styles, Barcelona, Catelonia*, 1991.

[9] G. Laan, "The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style," *Speech Communication*, vol. 22, pp. 43–65, 1997.

[10] R. Smiljanic and A. R. Bradlow, "Speaking and hearing clearly: Talker and listener factors in speaking style changes," *Language and Linguistic Compass*, vol. 3, no. 1, pp. 236–264, 2009.

[11] J.-P. Goldman, T. Prsir and A. Auchlin; Christodoulides, G., "Speaking style prosodic variation: an 8-hours 9-style corpus study," *Proceedings of Speech Prosody 7, Dublin, Ireland*, pp. 105–109, 2014.

[12] H. Levin, C. A. Schaffer and C. Snow, "The prosodic and paralinguistic features of reading and telling stories," *Language and Speech*, vol. 25, no. 1, pp. 43–54, 1982.

[13] J. Hirschberg, "A corpus-based approach to the study of speaking style," In: M. Horne (eds.). *Prosody: Theory and experiment*. Dordrecht/Boston/London. Kluwer, 2000, pp. 335–350.

[14] H. Mixdorff and H. R. Pfitzinger, "Analysing fundamental frequency contours and local speech rate in map task dialogs," *Speech Communication*, vol. 46, pp. 310–325, 2005.

[15] V. Dellwo, A. Leeman and M.-J. Kolly, "The recognition of read and spontaneous speech in local vernacular: The case of Zurich German," *Journal of Phonetics*, vol. 48, pp. 13–28, 2015.

[16] P. Wagner and A. Windmann, "Acted and spontaneous conversational prosody - same or different?," *Proceedings of Speech Prosody 8, Boston, USA*, 2016.

[17] H. Traunmüller and A. Eriksson, "Acoustic effects of variation in vocal effort by men, women, and children," *Journal of the Acoustical Society of America*, vol. 107, pp. 3438–3451, 2000.

[18] A. Eriksson and H. Traunmüller, "Perception of vocal effort and distance from the speaker on the basis of vowel utterances," *Perception and Psychophysics*, vol. 64, pp. 131–139, 2002.

[19] O. Niebuhr and L. S. Kohtz, "Eliciting speech with sentence lists – a critical evaluation with special emphasis on segmental anchoring," *Proceedings of Interspeech 2013, Lyon, France*, pp. 563–567, 2013.

[20] C. Cucchiarini, H. Strik and L. Bou, "Quantitative assessment of second lanugage learner's fluency: Comparison between read and spontaneous speech," *Journal of the Acoustical Society of America*, vol. 111, no. 6, pp. 2862–2873, 2002.

[21] M. M. North, S. M. North and J. R. Coble, "Virtual Reality Therapy: An effective treament for fear of public speaking," *International Journal of Virtual Reality*, vol. 3, no. 2, pp. 2–6, 1998.

[22] D.-P. Pertaub, M. Slater and C. Barker, "An experiment on public speaking anxiety in response to three different types of virtual audience," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 1, pp. 68–78, 2002.

[23] Presentation Simulator App, retrieved 11 November 2017 from http://www.presentationsimulator.com/fear-public-speaking/

[24] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer program]. Version 6.0.36," retrieved 11 November 2017 from http://www.praat.org/

[25] Y. Xu, "ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis," In *Proc. Tools and Resources for the Analysis of Speech Prosody (TRASP), Aix-en-Provence, France,* pp. 7-10, 2013

[26] P.A. Barbosa, "Semi-automatic and automatic tools for generating prosodic descriptors for prosody research," In *Proc. Tools and Resources for the Analysis of Speech Prosody (TRASP), Aix-en-Provence, France*, pp. 86-89, 2013.

[27] P. Touati, "Prosodic aspects of political rhetoric," *Proceedings of the ESCA Workshop on Prosody, Lund, Sweden*, pp. 168–171, 1993.

[28] A. Rosenberg and J. Hirschberg, "Acoustic/Prosodic and lexical correlates of charismatic speech," *Proceedings of Eurospeech 2005, Lisbon, Portugal*, pp. 1–4, 2005.

[29] A. Rosenberg and J. Hirschberg, "Charisma perception from text and speech," *Speech Communication*, vol. 51, pp. 640–655, 2009.

[30] O. Niebuhr, A. Brem and E. Novák-Tót, "Prosodic constructions of charisma in business speeches – A contrastive acoustic analysis of Steve Jobs and Mark Zuckerberg," *Proceedings of Speech Prosody 8, Boston, USA*, 2016.

[31] O. Niebuhr, J. Voße and A. Brem, "What makes a charismatic speaker? A computer-based acousticprosodic analysis of Steve Jobs tone of voice," *Compters in Human Behavior*, vol. 64, pp. 366–382, 2016.

[32] O. Niebuhr, S. Tegtmeier and A. Brem, "Advancing research and practice in entrepreneurship through speech analysis - from descriptive rhetorical terms to phonetically informed acoustic charisma metrics," *Journal of Speech Sciences*, vol. 6, no. 1, pp. 3–26, 2017.

[33] Y. Xu, "In defense of lab speech in prosody research," In G. Fant, H. Fujisaki and J. Shen (eds), Frontiers in Phonetics and Speech Science — Festtschrift for Professor Wu Zongji's 100th birthday. Beijing: The Commercial Press, pp. 381-390, 2009.

[34] M.M. North, S.M. North and J.R. Coble, "Virtual reality therapy: an effective treatment for the fear of public speaking," *The International Journal of Virtual Reality,* vol. 3, pp.1- 6, 1998.

[35] M Safir, P. Marilyn, H.S. Wallach and M. Bar-Zvi, "Virtual reality cognitive-behavior therapy for public speaking anxiety: one-year follow-up," *Behavior Modification,* vol. 36, pp. 235-246, 2012.

[36] V. Cunningham, M. Lefkoe, L. Sechrest, "Eliminating fears: an intervention that permanently eliminates the fear of public speaking," Clinical Psychology & Psychotherapy ,vol. 13, pp. 183-193, 2006.

[37] P.J. Denning and N. Dew, " The Profession of IT - The myth of the elevator pitch," *Communications of the ACM*, vol. 55, pp. 38-40, 2012.

[38] B. Schlender and R. Tetzeli, *Becoming Steve Jobs*. New York, Crown, 2015.