

Unguided VR public-speaking training enhances your confidence - but does not improve your intonation

Valls-Rates, I.; Niebuhr, Oliver; Prieto, Pilar

Published in:

Proc. 11th International Conference of Speech Prosody, Lisbon, Portugal

DOI:

10.21437/SpeechProsody.2022-33

Publication date:

2022

Document version:

Final published version

Document license:

Other

Citation for pulished version (APA):

Valls-Rates, I., Niebuhr, O., & Prieto, P. (2022). Unguided VR public-speaking training enhances your confidence - but does not improve your intonation. In *Proc. 11th International Conference of Speech Prosody, Lisbon, Portugal* (Vol. 2022-May, pp. 160-164). Article 33 <https://doi.org/10.21437/SpeechProsody.2022-33>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.

Unless otherwise specified it has been shared according to the terms for self-archiving.

If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk



Unguided VR public-speaking training enhances your confidence - but does not improve your intonation

İo Valls-Ratés¹, Oliver Niebuhr², Pilar Prieto^{3,1}

¹Universitat Pompeu Fabra

²Centre for Industrial Electronics, University of Southern Denmark, Sonderborg/DK

³ICREA

io.valls@upf.edu, olni@sdu.dk, pilar.prieto@upf.edu

Abstract

Public speaking is essential in our daily life. However, standing in front of a crowd is more often than not challenging for people. VR simulations can help speakers meet this challenge. Our study employed a between-subjects design with a VR group (N=17) and a Non-VR group (N=14). Both groups gave a 2-minute speech in front of a live audience before (PRE) and after (POST) they practiced public speaking in front of a VR audience or alone in a classroom (Non-VR). Each group had three of these VR or Non-VR training sessions, one per week. Acoustic analyses of both groups' PRE vs. POST prosodies show that 1) the two groups did not differ significantly in f0-related parameters as a function of training (f0 level, f0 range, f0 sd, f0 min/max), 2) the VR group has, unlike the Non-VR group, developed a stronger, clearer and more confident way of presenting, in terms of a longer talking time, fewer disfluent pauses, lower speaking rate, higher CPP and HNR, and lower jitter and shimmer levels. Thus, unguided VR training can help people give more persuasive speeches in real-life presentations, but we assume (consistent with previous research) that guided feedback is required to also improve people's speech melody.

Index Terms: VR, virtual, public speaking, anxiety, prosody.

1. Introduction

Boosting public speaking abilities in secondary school settings contributes to strengthening students' effectiveness with academic work, but also their social skills, thus affording them more satisfactory interpersonal relationships [1,2] and preventing them from abandoning their studies prematurely [3,4]. In order to achieve that, schools need to realize the importance that oral abilities have in the development of students' self-confidence. They obviously need to take action and consider the challenge of involving students in more oracy settings to enhance their communication, empathy and thus encouraging them in taking active part of the community they live in [1].

As any other skill, public speaking needs practice. [5] outline that speech practices are the single greatest predictor of student speaking success and key for reducing public speaking anxiety (PSA). Studies such as [6,7,8] show that self-assessed levels of communicative apprehension correlate negatively with self-perception of communicative competence. High levels of PSA can result in poor speech preparation [9] and decision-making of effective speech introduction strategies [10,11] and, most obviously, can negatively affect speaking performance [12,5]. The negative thinking of those speakers exhibiting larger levels of PSA can reduce their speaking competence [13] and impact their interpretation of feedback

[14]. Other research suggests that those speakers with low levels of PSA "may actually enjoy exhibiting their talents" [15] and are more self-assured about their speaking abilities [16,17]

The earlier speakers start preparing for their presentations, the more fluent (fewer total disfluencies) speeches will be [18]. However, practice needs time and it also needs to be oral. [19] reported that there is a high percentage of students who dedicate most of their time to writing their speech rather than to rehearsing it orally. In fact, students spend on average less than 5 minutes on rehearsing [19]. [20] found that, compared to students rehearsing alone, rehearsing in front of an audience gave students higher scores on their final classroom-speech assessment, thus lending support to the claim that audience-based speech practice can increase speech assessments.

As a way to enhance the oral practice of presentations and, also, to reduce anxiety when delivering speeches in front of an audience, VR simulations can be of great help. Practicing in virtual environments gives rise to the illusion of being in the environment depicted and make the user participate in it, experiencing the sense of presence [21,22]. Studies with university students show [23,24,25] how practicing with VR audiences results in diminishing communication apprehension and enhancing self-perceived communication competence. The study by [26] concludes that employing VR for speech rehearsals not only decreases PSA. Rather, students consider it an innovative way of oral rehearsing that makes them more willing to accomplish a good performance. [27,28] also refer to the excitement of students to participate in VR experiments as a different and motivating way to entice them to rehearse their speeches. The study of [28] compares three conditions for giving a single speech (in front of a mirror; using video recording, and using a VR setting). But, the VR setting did not yield higher scores in final classroom speech assessments than the other two conditions. The study also points out the fact that participants in the VR settings reported that it was more demanding than other modes of practice, which is consistent with the ability of digital audiences elicit stress in speakers.

Besides the fact that VR can provide a credible set of scenarios that allow for an immersive learning situation, when used for public speaking training VR environments also seem to be conducive to a more listener-oriented speaking style. [29] showed that participants rehearsing within a VR environment performed their speech in a more listener-oriented, conversation-like speaking style than participants in the control group, who practiced their speech alone in a classroom. They concluded that the speeches of participants who were trained in the VR condition were more charismatic and more audience-oriented, showing reduced signs of erosion due to repeated rehearsing than those speeches that had been practiced alone in a classroom. VR training also proved to be effective in

prompting vocal characteristics very similar to the ones used in the classroom in a very recent study [30]. It was conducted with elementary school teachers that gave the same lesson in their classrooms and later in front of a VR audience. Results showed how performing in front of real and virtual audiences (compared to free speech in a control condition) significantly increased the participants' f0 values, their f0 variations and their voice intensity (in line with [29]), but the VR setup had no effects on the number or duration of pauses.

Against the outlined research background, we see that still too little is known about the relationship between practicing oral speeches with VR and its effects on voice parameters and PSA. Therefore, we investigate, through a between-subjects training experiment, whether training with VR makes a difference in how the voice sounds and how the speaker feels in terms of self-perceived anxiety. To our knowledge this is the first experiment that investigates the effects of prosody in a post-training public speaking test with a live audience.

The following hypotheses are tested: (1) Compared to a baseline condition of non-VR speech training, VR-based speech training will, in a post-training condition in front of a real audience, result in less anxious speakers, i.e. in lower PSA values. (2) These lower PSA values will be reflected in prosodic differences compared to the baseline condition of speakers. (3) The more audience-oriented prosody triggered by VR-based speech training will carry over into the speakers' post-training speeches, unlike for the baseline-condition speakers. Note that, with reference to the above reviewed studies, we expect any form of training to have a (positive) effect on the post-training anxiety and public-speaking performance. The effects of VR-based training should differ from those of non-VR training and, e.g., be more strongly pronounced. That is, hypotheses (1)-(3) are, in statistical terms, essentially interaction hypotheses.

2. Method

2.1. Participants

A total of 31 secondary school students (mean age=16.95) participated in the study: the VR group (N=17) and the Non-VR group (N=14). Participation was voluntary, and all participants completed an informed consent process prior to enrolling in the study. Participants performed the speeches using Catalan. All students are bilingual Catalan-Spanish speakers, with a 89.7 % of self-reported dominance in Catalan.

2.2. Materials

The topic selected for both pre-training and post-training public speaking task was "Do you think that adolescents should spend more time in nature?". Before giving their speech, participants received a sheet of instructions in which they were asked to prepare and then deliver a two-minute speech in front of three representatives of the Catalan Department of Education to convince them to increase funding for secondary school field trips to the countryside. To stimulate the participants' thinking, five possible lines of argumentation were also provided. The instructions also noted that the participants would have two minutes in which to prepare their speech. Though they could take preliminary notes for that purpose if they wished, they would not be allowed to use the notes when they delivered their speech. The topics for each of the three trainings were the following: "What would the house of my dreams be like?", "Is graffiti a form of art?", and "Can happiness be bought?".

2.3. Training sessions

We conducted a between-subjects pre- and post-training experimental design with three training sessions, one per week. The total duration of the experiment, from pre-training to post-training was 5 weeks. Participants were allotted two minutes to prepare their speech and did so alone in an empty classroom. After the two minutes of preparation had elapsed, they went to the adjacent classroom. For the VR participants, the experimenter fitted them with a Clip Sonic® VR headset, which was attached to a smartphone. Using the free BeyondVR virtual reality interface application installed on the smartphone, the VR headset created the 3D illusion that the participant was standing in front of an audience. The virtual audience in this application moves while sitting and they show a sympathetic stance while the participant is speaking. They all look at the speaker and show interest in what the speaker is talking about. (Fig.1). Note that a timer is also visible in the view provided by the VR headset to allow speakers to monitor their use of time and not exceed the two-minute limit. For the Non-VR group, participants gave their speeches alone in a classroom and they could also check the timer placed in front of them. Participants of both groups were videotaped in all training sessions.



Figure 1: Screenshot of the VR scenario with a virtual audience generated by BeyondVR.

2.4. Anxiety measures

In order to control for anxiety and to facilitate comparisons with studies that have assessed anxiety in public speaking tasks through self-perception measures [31,24,32], we used the Subjective Units of Distress Scale, henceforth SUDS [33], a self-assessed anxiety measure which uses a 100-point scale anchored on 0 (no fear), 25 (mild fear), 50 (moderate fear), 75 (severe fear) and 100 (very severe fear). Participants were given the SUDS just prior to entering the room where they would give their pre and post-training speeches.

2.5. Procedure

All participants did the same pre-training, which consisted of giving a brief speech in front of a live audience (Fig.2).

The basic experimental structure of this between-subjects study was a pre-training followed by a training period and a post-training. The VR group performed the three training sessions delivering their speeches in front of a virtual audience, whereas the Non-VR group gave the speeches being alone in a classroom. The reason to choose three short VR sessions was based on the belief that adaptation to the that context would need some repetitions. Empirical reports of fast and reliable learning of visual context-target associations have proved effective after just three repetitions [34]). Finally, all participants carried out a post-training, which consisted of the

same task as the pre-training. Participants rated the degree of PSA they were experiencing right before each speech.

2.6. Data analysis: Acoustics and statistics

For each participant, the total durations of the recorded speeches were similar in the pre- and post-training conditions ($M = 1:34$ minutes; min-max = 1:15–2 minutes). A total of 62 speeches were analyzed, two (pre vs. post) for each of the 31 participants. Each analysis included 16 different parameters. The acoustic-phonetic analysis was automatically performed using the ProsodyPro script of [35] and the supplementary analysis script of [36], both with the (gender-specific) default settings of PRAAT [37]. In the pitch or fundamental-frequency (f_0) domain, we measured the minimum and maximum f_0 , the f_0 variability (in terms of the standard deviation), the mean f_0 and the f_0 range. For all five f_0 parameters, one value was determined per prosodic phrase. Measured values were checked manually for plausibility. Outliers or missing values were corrected by manual measurements. Moreover, all f_0 values were recalculated from Hz to semitones (st) relative to a base value of 100 Hz.

The duration/time domain consisted of the following four measured parameters: total number of silent pauses (> 300 ms, which is above the perceived disfluency threshold in continuous speech [39]), total time of the presentation (incl. silences), total speaking time (excl. silences), and net syllable rate (or articulation rate) in syl/s. All temporal measurements were conducted based on the analyzed presentation as a whole.

The domain of voice-quality measurements included the seven most measured parameters: harmonic-amplitude difference (h1-h2, f_0 corrected), cepstral peak prominence (cpp), harmonics-to-noise ratio (HNR), spectral center of gravity (CoG), formant dispersion (F1-F5), jitter, and shimmer. Like for the f_0 parameters, voice-quality measurements were conducted based on the prosodic phrase, i.e. one value per prosodic phrase was calculated. Also, all values were manually checked and corrected, if required.

All measurements were statistically analyzed by means of a mixed-model two-way MANOVA based on the two-level fixed factors Training (VR vs. non-VR) and Presentation (pre vs. post). Training was a between-subject and Presentation a within-subject factor. Participants were a random factor. Additional pairwise-comparisons tests used the Sidak correction to adjust the alpha-error level for multiple testing. Note that we are testing interaction hypotheses in the present paper. That is, we expect differences between the two levels of presentation, but, crucially, these differences are expected not to be the same for the VR and the non-VR groups of participants.



Figure 2: *The three-member audience attending pre-training and post-training sessions.*

The participants' SUDS ratings were also part of the MANOVA. SUDS values are no measurements, but the underlying scale can still be considered to meet the criteria of a parametric test procedure.

3. Results

The significant main effects of Training in the MANOVA concerned virtually all analyzed 16 parameters. Many of these main effects do not coincide with a significant Training*Presentation interaction. That is, what the MANOVA shows in these cases are offset differences between the two compared groups the VR group and the non-VR group. For example, for the participants in the VR group we generally measured higher f_0 values ($F[1,1932] > 7.86$, $p < 0.01$, $\eta_p^2 > 0.12$), including larger f_0 ranges ($F[1,1932] = 23.06$, $p < 0.01$, $\eta_p^2 > 0.18$), as well as higher h1-h2 values ($F[1,1932] = 25.55$, $p < 0.001$, $\eta_p^2 = 0.13$), lower CoG values ($F[1,1932] = 4.44$, $p = 0.04$, $\eta_p^2 = 0.05$) and fewer silent pauses ($F[1,1932] = 6.29$, $p = 0.02$, $\eta_p^2 = 0.10$) than for the participants of the non-VR group, independently of Training.

Regarding the main effects of Presentation, we also found a number of effects that occurred independently of VR or non-VR training, i.e. without an interaction effect. This includes, for example, a decrease in CoG ($F[1,1932] = 29.82$, $p < 0.001$, $\eta_p^2 = 0.15$) and an increase in HNR ($F[1,1932] = 23.01$, $p < 0.001$, $\eta_p^2 = 0.12$) from pre-training to post-training speeches. The f_0 variability also increased from pre- to post-training speeches after both VR and non-VR training alike ($F[1,1932] = 25.47$, $p < 0.001$, $\eta_p^2 = 0.22$). No other general (i.e. across-group) differences as a function of Presentation could be found.

The interaction effects we were looking for primarily did emerge for a number of parameters. None of these parameters concerned f_0 , though. In the duration/time domain, there was a significant Training*Presentation interaction for the total time of the presentation ($F[1,1932] = 5.23$, $p = 0.03$, $\eta_p^2 = 0.08$) as well as for the total speaking time ($F[1,1932] = 6.08$, $p = 0.02$, $\eta_p^2 = 0.09$) and the net syllable rate ($F[1,1932] = 5.89$, $p = 0.02$, $\eta_p^2 = 0.08$). According to the multiple comparisons tests, these interactions are all caused by a significant change in the VR group that is absent in the non-VR group. That is, unlike for non-VR training, the VR training made speakers present longer ($\phi +13$ s, VR = 117 s, non-VR = 104 s, $p < 0.05$) and speak more ($\phi +7$ s, VR = 71 s, non-VR = 65 s, $p < 0.05$) and slower ($\phi +0.3$ syl/s, VR = 5.5 syl/s, non-VR = 5.8 syl/s, $p < 0.01$) in their post-training speeches.

The interaction effects obtained for voice-quality are more complex in that we mostly see in the multiple comparisons tests significant changes from pre- to post-speeches for both the VR and the non-VR groups – but in opposite directions. For example, the VR groups showed a significant increase in cpp values, whereas those of the non-VR group decreased significantly from pre- to post-training speeches ($F[1,1932] = 5.71$, $p = 0.02$, $\eta_p^2 = 0.05$). Similarly, the jitter and shimmer values decreased significantly for the VR group, whereas they increased for the non-VR group from pre- to post-training speeches (jitter: $F[1,1932] = 13.44$, $p < 0.001$, $\eta_p^2 = 0.07$; shimmer: $F[1,1932] = 20.84$, $p < 0.001$, $\eta_p^2 = 0.13$). For the formant dispersion, the VR group showed a significant decrease from pre- to post-training speeches that was absent for the non-VR group ($F[1,1932] = 5.49$, $p = 0.02$, $\eta_p^2 = 0.04$).

Finally, as for the SUDS values we also see a significant main effect of Training due to an offset difference between the VR and the non-VR group ($F[1,1932] = 8.91$, $p = 0.004$, $\eta_p^2 = 0.13$). The latter group had on average a 13.0 points lower SUDS value ($\phi 40.5$) than the VR group ($\phi 53.5$). The effect of Presentation was not significant for SUDS, but, crucially, there was a significant interaction Training*Presentation ($F[1,1932]$

= 7.12, $p = 0.002$, $\eta_p^2 = 0.16$), based on the fact that the VR groups' SUDS values decreased significantly from pre- to post-training speeches by about 10.1 points (from $\bar{M} = 58.6$ to $\bar{M} = 48.5$, $p = 0.008$). By contrast, the non-VR groups' SUDS values did not decrease significantly but remained constant for the pre- and post-training speeches.

4. Discussion and Conclusions

The purpose of this experiment was to examine the impact of VR trainings on the secondary school participants' effective speeches when speaking in front of a live audience. In this way, we assess other ways of rehearsing speeches than doing it alone in a room. To achieve that, we designed a between subjects experiment with a pre-training, 3 training sessions and a post-training so that we could compare pre to post-training changes between VR and a baseline condition of Non-VR. The duration from pre to post-training was 5 weeks.

Acoustic analyses were performed across groups, and results showed that, compared to a baseline condition of non-VR speech training, VR training resulted in less anxious speakers, meaning that their SUDS values at post-training decreased by 10 points, whereas non-VR speakers remained at the same SUDS levels. These results go in line with previous studies where VR proved effective in reducing self-assessed PSA levels of participants [23-25,31]. This result is consistent with Hypothesis (1).

Regarding the effects of VR on prosodic parameters, results showed no significant changes across groups on f_0 values, meaning that intonation patterns did not change due to VR. At first glance this is inconsistent with the results of [30] where teachers performed the same lesson in class and with a virtual audience, or with the results of [29] where participants had to give persuasive investor pitches with and without a VR audience. The important difference to the present study is, however, that both [30] and [29] analyzed the prosody that speakers showed during VR immersion and not after it. As we already highlighted in the Introduction, our experiment is probably the first to analyze what happens (prosodically) when speakers take off the VR glasses again. In fact, as we will report in a different paper [38], the prosodic changes that we found during VR (and non-VR) training are largely consistent with both [30] and [29]. Insofar, one important implication of our study is that future studies have to examine in detail, what beneficial after-training changes VR training is able to trigger "automatically" in speech prosody, i.e. by the mere use of this technology, and what changes require an extra effort, for example, in the form of explicit feedback strategies. That f_0 -related melodic changes can basically be learned through training such that they last until speeches held after training was demonstrated by [40]. Unlike in our VR condition, the training condition of [40] included an explicit visualization and color-coded real-time evaluation of speech melody, though.

Regarding our Hypotheses (2) and (3), our results show that VR-based speech training made the participants' after-training speeches longer, more verbose, and slower in terms of the (net) tempo. We also found fewer silent pauses with durations > 300 ms that can make speakers sound disfluent in connected speech [39]. For voice-quality, we found that VR speakers increased their CPP and HNR levels and decreased their jitter and shimmer levels from pre-training to post-training speeches – whereas the exact opposite applied to the non-VR speaker sample. Overall, these effects and changes are consistent with the SUDS results in that they suggest the following: The speakers of the VR sample developed from pre- to post-training

a more comfortable and confident way of presenting as well as clearer, stronger and less "shaky" voice. The longer speaking times and higher HNR levels argue additionally in favor of more audience orientation. By contrast, the non-VR speakers developed in the opposite direction.

The latter, somewhat unexpected result is a by-product of the experimental design, but one that closely resembles real-life conditions. That is, the training that the non-VR speakers received, together with their awareness of having to present again in front of the same audience as in their pre-training speech created a high pressure to perform in the post-training session. The speakers had to assume that both the experimenter and their live audience would expect them to perform better than they did before the training; and unlike the VR Group, the speakers did not have any technology that could counteract this higher pressure through a PSA reduction. This explanation also fits in with the fact that the PSA level in the non-VR group did not decrease significantly. The gain in confidence, however, that the VR technology achieved probably based on the presence effect [21], persisted prosodically beyond the training, i.e. to the second speech the VR-trained speakers gave in front of their live audience. So, our findings should be taken as an encouragement for students to practice their speeches regularly and, for trainers, to spend more time on VR rehearsals in their courses.

In summary, this means that the present results provided empirical support for both Hypothesis (2) and (3). Yes, the VR sample's lower PSA values were notably reflected in prosodic differences compared to the non-VR sample; and yes, unlike for the latter sample, the VR sample's prosody is indicative of a more audience-oriented way of speaking that was triggered by the VR training and carried over into the post-training speeches.

The present study has some limitations. First of all, the size of the sample was relatively small: a larger sample would yield more robust results and, thus, a clearer picture of how VR training sessions affect 17-year-old's public speaking abilities. Also, in order to obtain effects on f_0 parameters, the study could have added explicit feedback strategies in order for VR to change or improve participants' intonation.

These limitations aside, we note that the majority of students enjoyed the VR sessions and found them valuable to face their upcoming oral projects, as other studies also re-marked [28,27]. Despite this, promoting realistic ways of individually rehearsing oral skills may enhance the whole experience of delivering a speech with regular and high-quality practice.

All in all, we believe that this study serves as a starting point to continue developing our knowledge about the relationships between VR practice, self-confidence and its effects on the speakers' prosody.

5. Acknowledgements

This work benefited from funding awarded by the Spanish Ministry of Economy and Competitiveness [PGC2018-097007-B-I00] and the Generalitat de Catalunya [2017 SGR_971]. We also acknowledge the support from the Recercaixa Project [RecerCaixa 2017ACUP_00249] and the Department of Translation at Universitat Pompeu Fabra. The 2nd author is the founder and CEO of the speech-technology company AllGood Speakers ApS. Please visit the following link for a conflict-of-interest statement: <https://www.allgoodspeakers.com/coi>

6. References

1. E. Bailey, "A Historical View of the Pedagogy of Public Speaking," *Voice and Speech Review*, 13(1), 31–42, 2018.

2. S.P. Morreale, M.M. Osborn, and J.C. Pearson, "Why communication is important: A rationale for the centrality of the study of communication," *Journal of the Association for Communication Administration*, 29(1), 125, 2000.
3. J. Boettcher, P. Carlbring, B. Renneberg, and T. Berger, "Internet-based interventions for social anxiety disorder - An overview," *Verhaltenstherapie*, 23(3), 160–168, 2013.
4. O. Niebuhr, "Advancing higher-education practice by analyzing and training students' vocal charisma: Evidence from a Danish field study," In *7th International Conference on Higher Education Advances (HEAd'21)*, Valencia, Spain, 743–751, 2021.
5. K.E. Menzel and L.J. Carrell, "The relationship between preparation and performance in public speaking," *Communication Education*, 43, 17–26, 1994.
6. K. Tekşan, H.H. Mutlu, and E. Çinpolat, "The examination of the relationship between the speech anxiety and speaking skill attitudes of middle school students and the opinions of teachers on speech anxiety," *Journal of Language and Linguistic Studies*, 15(4), 1395–1412, 2019.
7. J.W. Chesebro, J.C. McCroskey, D.F. Atwater, R.M. Bahrenfuss, G. Cawelti, J.L. Gaudino, and H. Hodges, "Communication apprehension and self-perceived communication competence of at-risk students," *Communication Education*, 41(4), 345–360, 1992.
8. P.R. Rosenfeld, R.A. Giacalone, and C.A. Riordan, "Impression management in organizations: Theory, measurement, and practice," New York: Routledge, 1995.
9. J.A. Daly, A.L. Vangelisti and D.J. Weber, "Speech anxiety affects how people prepare speeches: A protocol analysis of the preparation processes of speakers," *Communication Monographs*, 62, 383–397, 1995.
10. M.J. Beatty, "Public speaking apprehension, decision-making errors in the selection of speech introduction strategies and adherence to strategy," *Communication Education*, 37, 297–311, 1988b.
11. M.J. Beatty, and R.P. Clair, "Decision rule orientation and public speaking apprehension," In M. Booth-Butterfield (Ed.), *Communication, cognition, and anxiety*. [Special issue]. *Journal of Social Behavior and Personality*, 105–116, 1990.
12. M.J. Beatty, and R.R. Behnke, "Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance," *Human Communication Research*, 18, 147–176, 1991.
13. R.B. Rubin, A.M. Rubin, F.F. Jordan, "Effects of instruction on communication apprehension and communication competence," *Communication Education*, 46, 104–114, 1997.
14. M. Booth-Butterfield, "The interpretation of classroom performance feedback: An attributional approach," *Communication Education*, 38, 119–131, 1989.
15. M. McKinney, R. Gatchel and P. Paulus, "The Effects of audience size on high and low speech-anxious subjects during an actual speaking task," *Basic and Applied Social Psychology*, 4, 73–87, 1983.
16. J. Ayres and B.L. Heuett, "The relationship between visual imagery and public speaking apprehension," *Communication Reports*, 10, 87–94, 1997.
17. G.F. Thomas, W.G.Jr. Tymon, K.W. Thomas, "Communication apprehension, interpretive styles, preparation, and performance in oral briefing," *The Journal of Business Communication*, 31, 311–326, 1994.
18. A.M. Goberman, S. Hughes, T. Haydock, "Acoustic characteristics of public speaking: Anxiety and practice effects," *Speech Communication*, Volume 53, Issue 6, 867–876, 2011.
19. J.C. Pearson, J.T. Child and D.H.Jr. Kahl, "Preparation meeting opportunity: How do college students prepare for public speeches?," *Communication Quarterly*, 54, 351–366, 2006.
20. T.E. Smith and A.B. Frymier, "Get'real': Does practicing speeches before an audience improve performance?," *Communication Quarterly*, 54, 111–125, 2006.
21. M. Slater and M.V. Sanchez-Vives, "Enhancing Our Lives with Immersive Virtual Reality," *Frontiers Robotics and AI*, 3, 74, 2016.
22. K.C. Armel and V.S. Ramachandran, "Projecting sensations to external objects: evidence from skin conductance response," *Proceedings of the Royal Society of London. Series B*, 270, 1499–1506, 2003.
23. S.R. Harris, R.L. Kemmerling and M. North, "Brief virtual reality therapy for public speaking anxiety," *Cyberpsychology & Behavior*, 5, 543–550, 2002.
24. B.L. Heuett, and K.B. Heuett, "Virtual reality therapy: A means of reducing public speaking anxiety," *International Journal of Humanities and Social Science*, 1, 1–6, 2011.
25. S. Stupar-Rutenfrans, L. Ketelaars, and M.S. Van Gisbergen, "Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training in Home Environment Reduces Public Speaking Anxiety," *Cyberpsychology, Behavior, and Social Networking* 20(10), 624–633, 2017.
26. B.N. Frisby, R. Kaufmann, J.I. Vallade, T.K. Frey and J.C. Martin, "Using Virtual Reality for Speech Rehearsals: An Innovative Instructor Approach to Enhance Student Public Speaking Efficacy," *Basic Communication Course Annual*: Vol. 32, Article 6, 2020.
27. J.I., Vallade, R. Kaufmann, B.N. Frisby, and J.C. Martin, "Technology acceptance model: Investigating students' intentions toward adoption of 360° videos for public speaking rehearsals," *Communication Education*. Advanced online publication, 2020.
28. K. Kryston, H. Goble and A. Eden, "Incorporating virtual reality training in an introductory public speaking course," *Journal of Communication Pedagogy*, 4, 133–151, 2021.
29. O. Niebuhr and J. Michalsky, "Virtual reality simulations as a new tool for practicing presentations and refining public-speaking skills," In *Proceedings of the 9th International Conference on Speech Prosody 2018*, 309–313. *International Speech Communication Association (ISCA)*, 2018.
30. A. Remacle, S. Bouchard, A. Etienne, M. Rivard and D. Morsomme, "A virtual classroom can elicit teachers' speech characteristics: evidence from acoustic measurements during in vivo and in virtuo lessons, compared to a free speech control situation," *Virtual Reality*, 25(4), 2021.
31. D. Verano-Tacoronte and A. Bolívar-Cruz, "La confianza para hablar en público entre los estudiantes universitarios," *Conférence: XXIX AEDEM Annual Meeting*, June 2015.
32. C. Macinnis, S. Mackinnon, S., and P. Macintyre, "The illusion of transparency and normative beliefs about anxiety during public speaking," *Current Research in Social Psychology*, 15, 2010.
33. J. Wolpe, "Pergamon general psychology series, 1. The practice of behavior therapy," (4th ed.) Elmsford, NY, US: Pergamon Press, 1990.
34. M. Zellin, A. von Mühlénen, H.J. Müller, et al. Long-term adaptation to change in implicit contextual learning. *Psychon Bull Rev* 21, 1073–1079, 2014.
35. Y. Xu, "ProsodyPro - A Tool for Large-scale Systematic Prosody Analysis". In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, Aix-en-Provence, France. 7-10, 2013.
36. N.H. De Jong and T. Wempe, "Praat script to detect syllable nuclei and measure speech rate automatically," *Behavior research methods*, 41(2), 385–390, 2009.
37. P. Boersma and D. Weenink, "PRAAT: Doing Phonetics by Computer (Version 5.2.34)," [Computer software], 2007.
38. İ. Valls-Ratés, O. Niebuhr and P. Prieto, "Effects of public speaking virtual reality on prosodic and gestural features," *Proc. of Tone and Intonation in a globalized, digital world (TAI2021)*, Sønderborg, Denmark, to appear.
39. T. Lövgren and J.V. Doorn, "Influence of manipulation of short silent pause duration on speech fluency," *Proc. Disfluency in Spontaneous Speech Conference*, Aix-en-Provence, France, 123–126, 2005.
40. O. Niebuhr and J. Neitsch, "Digital Rhetoric 2.0: How to Train Charismatic Speaking with Speech-Melody Visualization Software," In: A. Karpov, R. Potapova (eds), *Lecture Notes in Computer Science* (Vol. 12335, Speech & Computer, 357–368). New York: Springer Nature, 2020.