# On the Convenience of Speeding Up Lecture Recordings: Increased Playback Speed Reduces Learning

Ida Ness[1], Kathinka Opdal[1], Frode Eika Sandnes[1,2][0000-0001-7781-748X]

[1] Oslo Metropolitan University, 0130 Oslo, Norway
[2] Kristiania University College, 0153 Oslo, Norway
ida.mn@outlook.com, s187533@oslomet.no, frodes@oslomet.no

**Abstract.** Audio and video lectures give students much freedom to study where and when they like. These digital media formats usually also give students flexibility to choose the playback speed. In this study we conducted a simple experiment to assess the effects of the playback speed on learning. A total of 20 participants were recruited using convenience sampling. The participants were divided into two groups, one half listening to a podcast at normal speed and the other at double speed, and thereafter completed a multiple-choice test. The results showed that the test scores were significantly lower at double playback speed. These results may have implications for teachers and students.

**Keywords:** playback speed, time-compression audio lectures, video lectures, recall, learning, accessibility, personalization

## 1 Introduction

Audio and specifically video-based lecture content has become widely available to students especially during the COVID-19 pandemic. Pre-recorded content gives students much freedom and flexibility in when they study, where they study and how they study. Another benefit of pre-recorded audio and video material is that students usually have the flexibility of adjusting the playback speed as well as going back and repeating specific segments [1]. Recordings are also believed to be beneficial for students with certain disabilities such as reduced vision [1] as one does not have to consider challenges associated with visual presentation [2] such as legibility of text [3, 4] but rather rely on non-visual cues [5, 6].

In a traditional physical lecture, students are bound to a time and place as well as following the pace of the lecturer. During course evaluations it is not uncommon for teachers to receive contradicting feedback indicating that the lectures are too slow, or too fast. Students who are unfamiliar with a topic or have a different language background may prefer a slow pace, while other students who already are familiar with the topic may want the lectures to move at a more rapid pace. Lecturers are therefore faced with the dilemma of diverging student needs. If they more slowly, all may comprehend the content, but one may risk losing the interest of the more advanced students. If one decides to move at a faster pace one risk losing the students that need a slower pace. With personal playback of pre-recorded contents, the students can select their preferred pace.

Much of the literature distinguishes between playback speed and time-compression or time-stretching [7]. Simply speeding up a recording will also increase the pitch and make spoken dialogue sound unnatural. This is the effect that can be heard if playing back a magnetic tape at a different speed than the original. Therefore, such applications typically employ some time-compression or time-stretching algorithm [8] so that the playback speed can be altered without affecting the pitch. For simplicity time-compression will be implied when referring to increased playback speed herein.

Informal dialogues with students reveal that some students indeed make use of this possibility, usually speeding up lectures that are too slow. There are also reports of an emerging trend where audiobook listeners listen to audiobooks at faster speeds. Users of screen readers with synthetic text-to-speech output typically speed up the speaking rate to cover more material per time unit. One may also imagine the opposite, that is, to decrease the playback speed when learning how to conduct certain tasks in detail, for instance when learning to play musical instruments or perform certain vocation-oriented tasks.

Playback speed is not the sole factor, also the actual speaking rate needs to be considered. A universal measure of speaking rate is words per minute (wpm), which is also used to measure text entry rates [9, 10, 11]. “Normal” speaking rate is believed to be around 150 words per minute, although individual variations can be large. In contrast, reading speeds are believed to be nearly twice that, namely 200-300 words per minute. One must also keep in mind that there are different types of reading such as skimming and careful (deep) reading. Obviously, the reading mode affects the reading speed.

We therefore wanted to explore to what degree an altered playback speed affects learning, in particular increased playback speed. We wanted to explore if playback speed can be comprehended when played at typical reading speeds. Our hypothesis was that increased playback speed will negatively affect learning. A controlled between groups experiment was designed where half of the participants listened to a podcast at normal speed and the other half listened to the same podcast at double speed upon which they completed a questionnaire. Although several similar experiments have been conducted in the past, we believe that it is important to replicate such experiments in the current technological context as the technology has advanced much since the early studies 50 years ago, both in terms of general accessibility, familiarity among end-users and audio quality.

This paper is organized as follows. The next section reviews related work. Section 3 presents the methodology. Results are presented in Section 4 followed by a discussion in Section 5. The paper is closed with concluding remarks in Section 6.

## 2 Related works

Several studies have explored the effects of speech rate on learning and some studies go back more than 50 years [12]. Barabasz [12] performed an experiment with more than 100 participants at graduate level by adjusting the playback speed of speech recorded on magnetic tape. The two conditions included normal speed and a second condition with the speed increased by 30%. The listening experiment was followed up with a test. The author found no effect of playback speed on recall.

King [13] conducted an experiment using time-compressed speech and found that the short-term listening performance started to deteriorate at 60% compression rate.

Ritzhaupt et al. [14] assessed performance with speech compressed at 40% and 80%. They found no effects on performance but did observe a significant effect on participants preference for the 40% compression rate.

Ritzhaupt and Barron [15] also conducted a large-scale study involving 300 participants with four conditions, namely normal playback speed, 50% increase, 100% increase and 150% increase. The results show a significant decrease in performance with 150% increase in playback speed compared to the other conditions.

Playback speed in conjunction with visual stimuli has also been investigated. A study involving 150 participants showed that increased playback speed combined with visual cues improved the recall but not the performance [16]. Goldhaber [17] found that both the playback speed and academic level of the listener were influential factors.

Speech rate has also been studied in context of screen reader users [18, 19]. Generally, experienced text-to-speech screen reader users configure the synthetic speech rate to twice, or three times, the normal rate, making it nearly intelligible for non-trained listeners. Also, screen-reader users often "skim-read" by skipping words, sentences and paragraphs using dedicated screen-reader controls. This allows screen reader users to cover more ground in a shorter time and hence more flexibility and freedom.

Speaking rate is also of interest to broadcasters [20]. For instance, for newscasts, it is important to employ a speaking rate that is comprehensible to the broad range of listeners which may have diverse backgrounds in terms of education and language proficiency.

## 3 Method

### 3.1 Experimental design

A between-groups experiment was chosen to ensure that each participant only was exposed to one of the two conditions. Playback speed was the independent variable with the two levels normal (original) speed and double speed. Success (recall) rate was the dependent variable.

### 3.2 Participants

A total of 20 participants were recruited using convenience sampling, of which 12 were female and 8 were male. All the participants were native Norwegian language users. Their mean age was 33.7 years of age (*SD* = 11.90). The youngest participant was 24 years old and the oldest 59 years old. A Mann-Whitney U test shows that the age of the males ($M$ = 35.3, $SD$ = 13.8) and the females ($M$ = 32.7, $SD$ = 11.1) were not significantly different ($W$ = 45, $p$ = .846). A non-parametric test was used as a Shapiro-Wilk test revealed that the age distributions were not normally distributed.

An attempt was made to divide the participants into two balanced groups using stratified sampling. The 20 participants were therefore split into two groups of 10 participants with 4 males and 6 females in each group. A Mann-Whitney U test confirmed that there were no significant differences in age across the two groups ($W$ = 58.5, $p$ = .542).

### 3.3 Material

For the listening experiment a professional Norwegian language podcast by Selda Ekiz published by the National Broadcasting Corporation in Norway (NRK) was used. The theme of this 8 minute and 50 second podcast is Why are fungi important to us? (Smartere på 10 minutter: Hvorfor er soppen viktig for oss?). This podcast was chosen as it was relatively short and thus suitable for the experiment. Moreover, the podcast is on a general topic which most participants could relate to. The podcast was professionally produced and therefore believed to be engaging to listeners.

Based on the content of the podcast a 24-question multiple choice test was designed with four alternatives for each question, in which only one option was correct. The questions addressed various information provided in the podcast. Google Forms was used to implement the questionnaire.

### 3.4 Procedure

The experiment was conducted remotely and electronically due to the COVID-19 pandemic. The participants were sent links to the podcast and questionnaire. It was not possible to adjust the playback speed on the podcast in the player on the website of the Norwegian Broadcasting Corporation. Therefore, the participants were asked to use the default podcast app available on their smartphones. This app is available on all smartphones and all the participants had smartphones.

Half of the participants played the podcast at its original speed according to its original duration of 8:50 minutes, while the other half of the participants listened to the podcast at double speed, namely 4:25 seconds. The participants who were assigned to listen to the podcast at double speed were instructed how to configure the podcast app.

The participants were asked to sit in a room without others present, and switch off any music, radio, television, or other potentially disturbing elements. They were instructed not to take notes or listen to the podcast, or parts of the podcast, several times, nor search for answers on the Internet. The experimenters pointed out that it was not their comprehension that was tested but the podcast.

The experiment was anonymous as it was conducted within a single session requiring no linking of observations [21].

### 3.5 Analysis

The success rate for each person was found by tallying all the correct answers and dividing these by the number of questions. The success rates for the participants were analyzed statistically using the open-source statistical analysis software JASP version 0.13.1.0 [22].

## 4 Results

Fig. 1 shows the result of the experiment. Clearly, those who listened to the podcast at normal speed achieved a nearly twice as high recall rate (*M* = 75.0%, *SD* = 13.0%) than those who listened to the podcast at double speed (*M* = 45.8%, *SD* = 10.8%). An independent sample t-test shows that this difference was statistically significant ($t(18) = 3.46$, $p < .001$, Cohen's $d = 2.441$). The large effect size confirms this difference.

We also assessed the participants for bias. A Welch t-test confirmed that there were no effects of gender ($t(10.270) = 0.354$, $p = .731$). A Welch test was used since a Levene's test revealed a deviation from the assumption of equal variances ($F(1) = 6.910$, $p = .017$). Moreover, age did not correlate with the correct rate ($r_s(20) = 0.021$, $p = .929$).

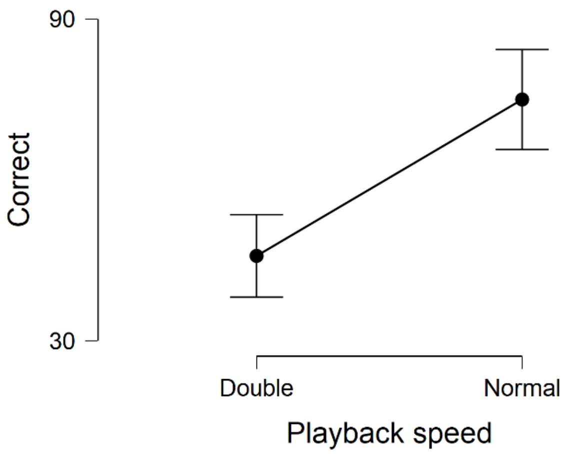


**Fig. 1.** Recall rates of normal and double playback speed (percentages). Error bars show 95% confidence intervals.

## 5 Discussion

The results support the hypothesis that reading speed affects recall, that is, if the audio playback speed is increased the recall is reduced. One possible explanation for this is that the audio played back at a higher speed is harder to comprehend. Another explanation is that the participants were given half the time to absorb the material. Perhaps the normal playback rate gives the participants sufficient space to process and briefly reflect over the information provided, while at half speed the participants were unable to do much apart from listening. On the other hand, there was obviously a longer time delay between the presentation of the information and the time to give the response for the participants who listened at normal speed, which could also have contributed to participants forgetting some of the contents, while those who listened to the podcast in half the time had a comparatively smaller chance of forgetting the content. The experiment was not designed to distinguish between the root cause, whether it is the legibility of the audio or other factors that results in the lower recall rates.

The results agree with previous findings such as results reported by King [13] and Ritzhaupt and Barron [15], while they disagree with the results of other studies such as Barabasz [12], Ritzhaupt et al. [14]. One possible explanation for the diverging results could be the quality of the audio technology used in the studies, as well as the degree of playback speed change. Clearly, studies with a large change in playback speed are probably more likely to trigger effects on recall than small playback speed changes. However, one should be careful when contrasting the results from these different studies since the experimental setups were not standardized.

## 6 Limitations

There are several weaknesses with the current experiment. First, the sample was relatively small with a comparatively wide spread of participants. It would be relevant to repeat the experiment with a larger sample and perhaps more focused cohorts of pupils (up to and including secondary school level) and students (at university level).

We did not probe the background of the participants and in this instance participants with a background in biology would have found the questionnaire easier than those without explicit training in biology. In hindsight we should have also asked about this, even with narrow cohorts, as participants may have diverse backgrounds with a myriad of study combinations or personal interests and hobbies. Moreover, we did not probe if any of the participants had reduced cognitive function such as reduced memory, as this could also have impacted the recall rates [23]. We should also have asked participants if they had any hearing.

The experiment was conducted remotely, and we were therefore unable to control the environment in which the experiment took place. It is therefore possible that some participants had background noise or other disturbances that may have affected the results [24, 25, 26].

Another aspect which was not addressed herein was the difficulty of the message. As the readability of written language affects comprehension [27, 28, 29], it is also possible that the language used in the podcast could have impacted the results. One way to check for this would be to assess the theoretical readability of the content by applying readability indices to a transcript of the podcast contents. However, the podcast was produced by the Norwegian natural broadcaster and presumably has gone through certain quality assurance procedures. Moreover, all the participants listened to the same content and there would thus be no bias across the two groups.

In hindsight it would have been relevant to probe the participants preferred playback speed. Preferred playback speed is highly personal and depends on factors such as listening context and training in listening to fast audio.

## 7 Conclusion

A study exploring audio lecture playback speed on learning was reported. The results showed that audio content played back at double speed led to reduced recall rate during a multiple-choice post-test. The results thus confirm our hypothesis that playback speed affects learning. These results may have implications for teachers who deliver pre-recorded audio or video lecture material. In this between groups experiment the participants had to listen to the audio in the designated speed category. In a real pedagogical setting the student will have a choice of speeding the audio or not, with the default being normal playback speed. Students who are more comfortable with speeding up the playback speed may do so permanently or during parts of a recording. However, it is likely that this may compromise comprehension and the students should be made aware of this. On the other hand, speeding up the content may help retain students' interest in the topic. It may be worthwhile to compromise the comprehension of some details, as it is probably better to have some learning outcome than no learning outcome. In any case, the freedom for students to choose their preferred playback speed is most likely a benefit to the learning process. Teachers should take explicit steps to ensure that students are aware of such functionality. This study did not address the degree with which students are aware of playback speed controls. Future work could explore students' perceptions of, and experience with, such speed control mechanisms. Another avenue of exploration is playback speed of videos where audio stimuli are accompanied by visual stimuli, both in textual format and as teachers' facial cues and gesticulations [30].

### References


1. Junor, L.: Teaching by tape: Some benefits, problems, and solutions. Distance Education **13**(1), 93-107 (1992). https://doi.org/10.1080/0158791920130108
2. Sandnes, F. E.: What do low-vision users really want from smart glasses? Faces, text and perhaps no glasses at all. In: International Conference on Computers Helping People with Special Needs, pp. 187-194. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41264-1_25
3. Sandnes, F. E., Zhao, A.: An interactive color picker that ensures WCAG2. 0 compliant color contrast levels. Procedia Computer Science **67**, 87-94 (2015). https://doi.org/10.1016/j.procs.2015.09.252
4. Sandnes, F. E.: Understanding WCAG2.0 color contrast requirements through 3D color space visualization. Stud. Health Technol. Inform, **229**, 366-375 (2016). https://doi.org/10.3233/978-1-61499-684-2-366
5. dos Santos, A. D. P., Medola, F. O., Cinelli, M. J., Ramirez, A. R. G., Sandnes, F. E.: Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants. Universal Access in the Information Society **20**, 93–103 (2021). https://doi.org/10.1007/s10209-020-00712-z

6. Lin, M. W., Cheng, Y. M., Yu, W., Sandnes, F. E.: Investigation into the feasibility of using tactons to provide navigation cues in pedestrian situations. In: Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, pp. 299-302 (2008). https://doi.org/10.1145/1517744.1517794
7. Pastore, R., Ritzhaupt, A. D.: Using time-compression to make multimedia learning more efficient: Current research and practice. TechTrends **59**(2), 66-74 (2015). https://doi.org/10.1007/s11528-015-0841-2
8. Omoigui, N., He, L., Gupta, A., Grudin, J., Sanocki, E.: Time-compression: Systems concerns, usage, and benefits. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 136-143. ACM (1999). https://doi.org/10.1145/302979.303017
9. Sandnes, F. E., Thorkildssen, H. W., Arvei, A., Buverad, J. O.: Techniques for fast and easy mobile text-entry with three-keys. In: Proceedings of the 37th Annual Hawaii International Conference on System Sciences 2004. IEEE (2004). https://doi.org/10.1109/HICSS.2004.1265675
10. Sandnes, F. E.: Evaluating mobile text entry strategies with finite state automata. In: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, pp. 115-121. ACM (2005). https://doi.org/10.1145/1085777.1085797
11. Aschim, T. B., Gjerstad, J. L., Lien, L. V., Tahsin, R., Sandnes, F. E.: Are split tablet keyboards better? A study of soft keyboard layout and hand posture. In: IFIP Conference on Human-Computer Interaction, pp. 647-655. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29387-1_37
12. Barabasz, A. F.: A study of recall and retention of accelerated lecture presentation. Journal of Communication **18**(3), 283-287 (1968). https://doi.org/10.1111/j.1460-2466.1968.tb00077.x
13. King, P. E., Behnke, R. R.: The effect of time-compressed speech on comprehensive, interpretive, and short-term listening. Human Communication Research **15**(3), 428-443 (1989). https://doi.org/10.1111/j.1468-2958.1989.tb00192.x
14. Ritzhaupt, A. D., Gomes, N. D., Barron, A. E.: The effects of time-compressed audio and verbal redundancy on learner performance and satisfaction. Computers in human behavior **24**(5), 2434-2445 (2008). https://doi.org/10.1016/j.chb.2008.02.017
15. Ritzhaupt, A. D., Barron, A.: Effects of time-compressed narration and representational adjunct images on cued-recall, content recognition, and learner satisfaction. Journal of Educational Computing Research **39**(2), 161-184 (2008). https://doi.org/10.2190%2FEC.39.2.d
16. Ritzhaupt, A. D., Barron, A. E., Kealy, W. A.: Conjoint processing of time-compressed narration in multimedia instruction: The effects on recall, but not recognition. Journal of Educational Computing Research **44**(2), 203-217 (2011). https://doi.org/10.2190%2FEC.44.2.d
17. Goldhaber, G. M.: Listener comprehension of compressed speech as a function of the academic grade level of the subjects. Journal of Communication **20**(2), 167-173 (1970). https://doi.org/10.1111/j.1460-2466.1970.tb00874.x
18. Borodin, Y., Bigham, J. P., Dausch, G., Ramakrishnan, I. V.: More than meets the eye: a survey of screen-reader browsing strategies. In: Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility. ACM (2010). https://doi.org/10.1145/1805986.1806005
19. Lee, H., Jang, Y., Hong, K. H.: A Comparison of the Listening Speed of the Korean TTS For the Blind: Based on their Screen Reader Experiences. In: International Conference on Computers for Handicapped Persons, pp. 49-52. Springer, Cham (2014).
20. Smith, J. R., McEwen, W. J.: Effect of newscast delivery rate on recall and judgment of sources. Journal of Broadcasting & Electronic Media **18**(1), 73-78 (1973). https://doi.org/10.1080/08838157309363721

21. Sandnes, F.E.: HIDE: Short IDs for Robust and Anonymous Linking of Users Across Multiple Sessions in Small HCI Experiments. CHI '21 Conference on Human Factors in Computing Systems Extended Abstracts Proceedings. ACM, (2021). https://doi.org/10.1145/3411763.3451794
22. JASP Team (2020). JASP (Version 0.14.1)[Computer software].
23. Sandnes, F. E., Lundh, M. V.: Calendars for individuals with cognitive disabilities: a comparison of table view and list view. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, pp. 329-330. ACM (2015). https://doi.org/10.1145/2700648.2811363
24. Vasilev, M. R., Kirkby, J. A., Angele, B.: Auditory distraction during reading: A Bayesian meta-analysis of a continuing controversy. Perspectives on Psychological Science **13**(5), 567-597 (2018). https://doi.org/10.1177%2F1745691617747398
25. Thompson, W. F., Schellenberg, E. G., Letnic, A. K.: Fast and loud background music disrupts reading comprehension. Psychology of Music **40**(6), 700-708. (2012). https://doi.org/10.1177%2F0305735611400173
26. Halin, N.: Distracted while reading? Changing to a hard-to-read font shields against the effects of environmental noise and speech on text memory. Frontiers in psychology **7**, 1196 (2016). https://doi.org/10.3389/fpsyg.2016.01196
27. Kaushik, H. M., Eika, E., Sandnes, F. E.: Towards Universal Accessibility on the Web: Do Grammar Checking Tools Improve Text Readability?. In: International Conference on Human-Computer Interaction, pp. 272-288. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49282-3_19
28. Eika, E., Sandnes, F. E.: Authoring WCAG2. 0-compliant texts for the web through text readability visualization. In: International Conference on Universal Access in Human-Computer Interaction, pp. 49-58. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40250-5_5
29. Eika, E., Sandnes, F. E.: Assessing the Reading Level of Web Texts for WCAG2. 0 Compliance—Can It Be Done Automatically?. In: Advances in Design for Inclusion, pp. 361-371. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41962-6_32
30. Sandnes, F. E.: Absence and Presence of Faces in Videos during the COVID19 Lockdown. In 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, pp. 189-193. ACM (2020). https://doi.org/10.1145/3439231.3439235t