Instructions for Staging and Recording a Lipreading-Deaf-Accessible Video of a Lipreading-Deaf-Accessible Zoom Lecture in a Manner that Improves Accessibility to the Blind

Daniel M. Berry

David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1, Canada
dberry@uwaterloo.ca
7 July 2021

1 Introduction

These days, many meetings are being held on e-meeting platforms. Generally, during a meeting, the platform (1) shows images of the participants speaking, (2) delivers the sounds of their speaking and of their backgrounds, and (3) shows images of documents the participants share. As these meetings are usually done, they are not particularly accessible to the deaf and the blind. While the sound quality is usually more than acceptable to the hearing, the deaf cannot hear it. No small-sized video of anyone's speaking is lipreadable. Even a large-sized video of the active speaker is not lipreadable because its image is not refreshed often enough for lipreading. The video image of a shared, even textual, document, can be read neither by the blind nor by their reading software. Video recordings of these meetings suffer the same problems, and it's even worse because when the recording is being played, (1) if a speaker's head has been recorded too small to lipread, or (2) slides have been recorded too small too read, nothing can be done to make it otherwise. Moreover, no one can ask a participant what E[1] said or request that what was said be written into a chat window. The recording prepared by a speaker as a backup in case she[2] cannot connect to a meeting at the time that her lecture is scheduled in the meeting can suffer the same problems, and the playing of this video during the live meeting will not be accessible.

This document attempts to ameliorate these problems by describing how

- to stage a lecture with slides given by a speaker on the Zoom[3] e-meeting platform, in a manner that is accessible to the lipreading deaf, and
- to record this lecture in a manner that makes a video that is accessible to the lipreading deaf.

[1] When this document needs a *singular* personal third-person pronoun for a person of unknown gender, it uses "E", "em", and "er" in the subjective, objective, and possessive modes, respectively. It eschews "they", "them", and "their" as being plural and thus introducing number and plural ambiguities.

[2] Without loss of generality, here and in the rest of this document, (1) a speaker is a she; (2) any of a Zoom host doing a recording, a moderator introducing a speaker, a blind or deaf person attending the live lecture, or a blind or deaf person watching and listening to the video recording of the lecture is a he; and (3) any other person, e.g., attending a live lecture and asking questions, is an E. Thus, since there is only one of the kinds of person that is a he in most discourses, pronouns are usually disambiguated.

[3] Why Zoom and not other platforms?
Note that I do *not* own stock in Zoom, Inc. ☺. It's just that, to date, there is no e-meeting platform that I know of whose speaker images are as lipreadable as is Zoom. Even with a live speaker whose computer is connected by a high-speed *wired* connection, even within a university's own high-speed VPN, neither Discord, Kast, Meeting, Skype-for-Business, Teams, nor WebEx manages to refresh the speaker's image rapidly enough to permit lipreading, even though the speaker's image may appear to move smoothly to the untrained — in lipreading — eye. With all that is happening during a live meeting, with all the media that the platform must keep up to date, sometimes at video refresh rates, the surprise is *not* that a platform does not refresh the speaker's video rapidly enough for lipreading. The surprise is that Zoom usually *does* refresh the speaker's video rapidly enough for lipreading. It seems that Zoom's ability to refresh the speaker's image in speaker mode rapidly enough, no matter the number of participants, is Zoom, Inc.'s competitive advantage.

As a side effect, following the instructions improves accessibility of the slides to the blind in both the lecture and in the video.

Sections 1.1 and 1.2 amount to an executive summary of the whole rest of this document. If you are familiar enough with the operation of Zoom that you can see exactly what to do from this executive summary, then by all means, don't waste your time reading the rest of this document, and just do it!

1.1 What

The key idea shared by the solution to both problems is that

- the image of the speaker's head can be or is made as large, and thus as lipreadable, as possible in the Zoom window, given the circumstances, at the cost of
- making the image of the shared slides no larger than is needed in the circumstances, which might include that
- the actual shared slides, viewable with their original application, are available to be viewed outside the Zoom window, via this application, by each meeting participant or video viewer. The actual slides, viewed with their original application, can be read by a text-to-voice reader by a blind participant or viewer.

In the case of a live meeting, each participant decides how much of er Zoom window is to be allocated to the image of the speaker's head, and in the case of a recorded video, the video is recorded in order to make the speaker's head is made as large possible given whether or not the actual shared slides are available to the video viewer.

The basic idea for a recording is to offer to the viewers of the resulting video

- the video itself, and
- a copy of the actual shared slides, viewable with their original application. This copy of the slides, viewed with its original application, can be read by a text-to-voice reader by any blind viewer

1.2 How

The recording of the lecture must be done in a non-standard way.

- The recording must be done on a machine, i.e., the host's, *other than* that from which the speaker's slides are shared, which is usually the speaker's machine.
- The host, who does the recording must adjust the division of the Zoom window into
  - the image of the speaker's shared slides and
  - the image of the speaker's head

  in a manner that
  - maximizes the size of the image of the speaker's head while
  - keeping the size of the image of the shared slides just big enough for the viewer to know where in his local copy of the speaker's slides he should be.

  This division is then recorded.

If the speaker has *not* shared her slides with the participants, then the division of the zoom should be about 50–50, because the participants will need to be able to *actually read* the image of shared slides.

Of course, the devil is in the details.

1.3 Plan for Sections

The instructions are quite detailed and might seem totally nonsensical. To motivate the hairy details so that (1) they actually make sense and are easy to remember and (2) they are easily adapted when the Zoom platform undergoes its inevitable changes in the future, the document describes (1) in Section 2, what a deaf person and a blind person do during a live lecture in a Zoom meeting, and (2) in Section 3, what happens when a slide sharer records a Zoom meeting. Section 1.6 lists assumptions about blind and deaf people that are assumed in the subsequent instructions. Then, Section 4 gives detailed instructions for making a video recording of a Zoomed lecture that is accessible to the deaf, while not excluding the blind. These instructions include steps for turning on Zoom's closed captioning,

a.k.a., live transcript, to further improve accessibility to the deaf. Section 5 describes how these instructions need to be modified for a speaker to record her own lecture in a manner that is accessible to the deaf, while not excluding the blind. The instructions require cooperation on the speaker's part. Therefore, Section 6 provides a sample letter to be sent to a speaker that describes what she must do to help make any recording of her accessible to the deaf, while not excluding the blind. Finally, Section 7 gives pointers to information about adding high quality closed captioning to a completed video recording of a lecture. An appendix lists the full recording instructions, as specified in Section 5, for self-recording. The online appendix directory at `https://cs.uwaterloo.ca/~dberry/FTP_SITE/tech.reports/RecordingInstructions/` gives some additional information, including these instructions and two examples of well-done recordings of fun lectures by Daniel Berry. One example has the shared slides minimized and the speaker's face maximized, and the other example has the shared slides and the speaker's face balanced. Edited captions have been added to these videos. See the README files in the directory and in the directory's subdirectories for further details about the contents of this appendix.

If you are interested in instructions only for staging a Zoom lecture in a manner that is accessible for the lipreading deaf without excluding the blind, read only

1. the rest of Section 1,
2. Section 2,
3. Section 4 until the end of Step 3 of the procedure,
4. and then Step 7 until the end of Substep (d), and
5. Section 6.

1.4 About the Screenshots

I captured screenshots that *show* the instructions being followed[4]. The meetings from which screenshots were taken consisted of two participants, (1) Daniel Berry as host and (2) Daniel Berry on another computer where the speaker Dubi Bear is speaking — albeit silently — using Daniel Berry's account. In*dubi*tably, there was no need to obtain the speaker's permission to be the subject of these screenshots ☺. The speaker instance of Daniel Berry has the same rights and powers in a Zoom meeting as any other participant, except as specifically assigned to this participant by the host.

1.5 Caveats and Disclaimers

A few caveats are in order. I, the author, am nearly deaf since birth. I base what I write in these instructions on my personal experience, which may or may not be representative of that of all hearing impaired people. I supervised a blind PhD student and learned how I had to interact with my student by e-documents and e-meeting platforms.

I did the best I could to validate everything written here by experimenting with various options. In the end, I had to generalize from few instances all done on two Macbook computers running Mac OS X Catalina, running Zoom 5.2.2 for Macs, logged in to my Zoom Pro account. I could not find definitive documentation at `support.zoom.us` for a lot that is claimed to be true here. I did my best to accurately describe the buttons, messages, and other widgets from the Zoom window's user interface, but I may have mistranscribed things. Worse, I may have tacitly assumed knowledge on your part.

In particular, I have seen evidence that the Zoom clients on different computers and operating systems behave differently enough that the instructions here do not work in all situations. For example, someone using the Linux Zoom client was not able to reproduce the screen layout described in these instructions. Also, a host running on the Linux Zoom client is not able to enable AI-generated live captioning, because there is no button labeled "CC Live Transcript" among the buttons at the bottom of his Zoom window.

All this being said, *please* do a dry run with a local colleague or student giving a lecture *before* you attempt to apply these instructions to record a live lecture by a guest from the outside. Also, if you are attempting to replace an existing lecture recording procedure with that described in these instructions, *please* continue to do the recordings with

[4] These snapshots were not captured in the *one* Zoom meeting, mainly because I realized only after finishing one meeting that I needed some additional screenshots. Sigh! Sometimes a screenshot is used for multiple purposes. Therefore, the visual continuity of screenshots that should be sequential is not guaranteed.

the old procedure in *parallel* with doing the recordings with the new procedure, *until* you are completely comfortable with the new procedure.

Finally, please contact me at dberry@uwaterloo.ca to report any observed errors with suggested corrections or to report better ideas for achieving the same goals, even on platforms other than Zoom.

1.6 Background Assumptions

Understand that many a blind person uses software that can read lecture slides in a so-called *machine-readable form* and say, with an artificially generated voice, what the slides say. An artifact is in machine-readable form if and only if it is a file created by an application that stores the word contents of the artifact in some explicit coding of the characters of the artifact's text. Examples of machine-readable forms include (1) .ppt or .pptx files, produced, e.g., by PowerPoint, and (2) .pdf files, produced, e.g., by Acrobat's Distiller, that contain the characters and not bitmaps of the text. A bitmapped image of text, such as (1) is produced by scanning[5] or (2) is displayed on a screen by a slide-showing application, is *not* in machine-readable form. Thus, this reading software cannot read any video's bitmapped image of the slides.

Understand that many a deaf person needs to read the speaker's lips to understand her and will want the image of the head to be as large as possible within his Zoom window. Making the speaker's head large generally requires making the image of the speaker's shared slides small and vice versa, because there is only so much space in a video window. The general strategy adopted in these instructions is that the video will minimize the slides to a thumbnail no larger than necessary for the viewer to be able to know where he should be in a local copy of the speaker's slides, viewed in its own application. The speaker's head will then be maximized to fill the rest of his Zoom window.

Therefore, it is assumed that a copy of the speaker's slides in machine-readable form is available for viewing or downloading along with the video recording of the lecture[6]. It is useful for these slides to be available also for the benefit of at least any blind or deaf attendee of the live lecture on an e-meeting platform.

Also, the speaker must be persuaded *not* to speak from a smartphone and to use a computer. Apparently, a Zoom app that run on a smartphone is incapable of transmitting both the contents of the camera, focused on the speaker's face, and the image of the shared slides at the same time. It transmits either one or the other. This feature of smartphone Zoom apps is a disaster from the lipreading participant's viewpoint.

Many a deaf person reads lips. Independently, many a deaf person uses some sign language, such as North American Sign Language (NASL), and the sign language is his preferred means of communication. Many deaf both read lips and sign. Some do only one, and of the deaf that do only one, a majority sign rather than read lips. Therefore, if you wish to be totally accessible to signing-only deaf, then you must engage a professional simultaneous signer to sign as the speaker is speaking and to record into the video a large enough image of the signer signing. However, a simultaneous signer is expensive, *and* probably all of the deaf that will be interested in watching the video also read lips, text, captions, and slides. So, the video's being accessible to the lipreading deaf is probably sufficient. Consequently, in the rest of this document "deaf" means "lipreading deaf".

In any case, any deaf person that might conceivably be watching the video of a lecture is able to read text at the university level. Therefore, synchronized closed captions[7] that are a verbatim transcript of what the speaker in a video says guarantees reasonable accessibility of the video to the deaf. In order to be useful, captions must be at least phonetically correct. A deaf person can deal with incorrect spelling, incorrect or missing punctuation, incorrect sentence division, and even incorrect word division if what he reads is phonetically identical to what a hearing person hears. The same human intelligence that correctly understands speech can correctly understand such phonetically correct captioning.

To get really good useful synchronized captions requires a human, with the skill and speed of courtroom stenographer, transcribing in real time what is being said. Such a human being is probably too expensive for the kinds of venues these instruction will be used. So, in most cases, the captioning will be automated, AI-generated live captioning

[5] unless the scanned image has been processed and modified by, e.g., Acrobat's text recognizer, to be searchable and selectable

[6] If the speaker is reluctant to make the slides available in an editable form, e.g., as a vanilla .ppt or .pptx file, she should be encouraged (1) to make the file read-only or modifiable only with a password or (2) to export it to an easily locked form such as a .pdf file.

[7] *Closed captioning (CC)* is a system that provides subtitles for television programs, accessible through a decoder (online Dictionary Version 2.3.0, on Apple's Mac OS X Catatlina)

provided by the meeting platform, such as Zoom. However, no automated, AI-generated live captioning comes close to being perfectly accurate. In fact, a skilled lipreader, using his real intelligence (RI) is significantly more accurate at understanding what a speaker is saying than any AI-generated live captioner. This fact is why achieving a lipreadable refresh rate for the speaker's video is so important, *even* when there is good AI-generated live captioning available.

Synchronized captions are not a substitute for Section 508-Compliant[8] closed captioning inserted into a video recording of a completed lecture by a human, domain-savvy transcriber, as described in Section 7. However, any kind of synchronized captioning, even AI-generated live captions, should help a lipreader tide through periods in which the speaker's video is not refreshed frequently enough for reliable lipreading, particularly if at the same time the quality of the sound has not also deteriorated. See Section 7 for a fuller discussion about the usefulness of closed captioning to a lipreader.

2 During A Live Lecture on Zoom

It is useful to understand how a deaf person operates the Zoom session with which he attends a live lecture.

On any device, a person can have at most one Zoom meeting active at any time. Therefore, the deaf person can have only one full-sized Zoom window on his device, which can be a computer, a tablet, or even a cell phone. Thus, the shared slides and the speaker's head will have to be distributed over one Zoom window that, at best, covers the entire screen of the device. Since a recording has to record both the slides and the speaker's head in one video, the Zoom session on which the recording is done is under the same constraint.

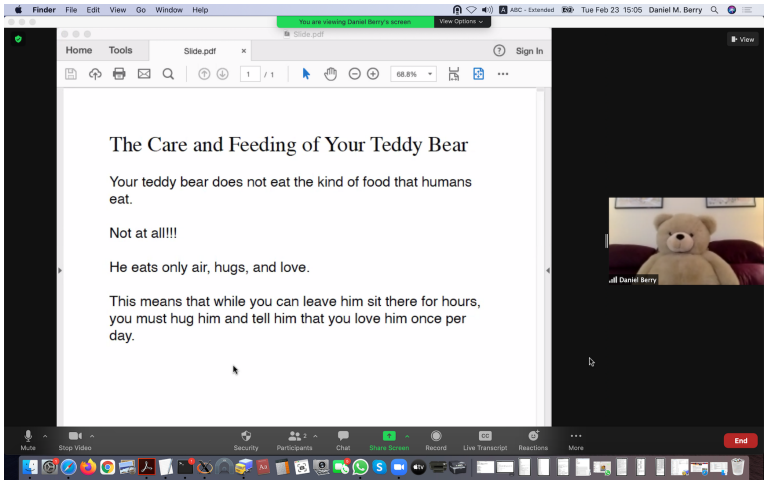

Fig. 1. Initial Division of Zoom Window After Sharing While in Speaker Mode

As the lecture begins, the speaker has shared her slides. The deaf person chooses the "Speaker View" option[9]. As shown in Figure 1, there is a vertical border, consisting of two short thin parallel lines, between (1) the shared slides, on the left side of the Zoom window, and (2) the speaker's head, on the right side of the Zoom window. The border can

[8] See `https://www.section508.gov/create/video-social`.

[9] The capability to choose "Speaker View" and to control the division of the Zoom window as described in this section is not available to ordinary participants in a webinar Zoom meeting. If a deaf person finds himself in a webinar Zoom meeting, he should unmute himself and request to be promoted to being a panelist in exchange for stopping his video and muting. If necessary, he should explain that as a webinar participant, he is not able to control the relative sizes of the speaker's and the shared slides' images in order to allow both reading the speaker's lips and reading the speaker's slides; he needs to be a panelist to have the usual full control of a meeting participant. If the host does not know how to promote, the deaf person should explain, "Click on the 'Participants' button at the bottom of you Zoom window; in the resulting Participants window, hover over my name until 'More' shows up; click on 'More'; and then click on 'Promote to Panelist'. Thank you."

be moved to the right to maximize the slides while minimizing the speaker's head, as is shown in Figure 1 itself. The border can be moved to the left to minimize the slides while maximizing the speaker's head, as is shown in Figure 11. Figure 12 shows yet another division.

That the current view is speaker mode in each of these figures is clear from the "View" icon in the upper right hand corner of the Zoom window. The symbol in the icon represents the big slide image adjacent to the small speaker image. In gallery mode, the single small dot on the right hand side of the symbol is replaced by three small dots stacked vertically to the right of the big slide image representative.

The way the deaf person operates his Zoom window depends on whether he has a local copy of the speaker's slides.

- If he has a local copy of the speaker's slides, then at the beginning of the lecture, he moves the border to the left to minimize the slides to a thumbnail no larger than necessary to be able to know where he should be in the local copy of the speaker's slides. The speaker's head is as large as possible. He listens to the speaker, reading her lips and following the slides with his local copy of the slides, keeping it synchronized to the current state of the thumbnail.
- If he does not have a local copy of the speaker's slides, then whenever during the lecture, a new slide comes up,
  1. he first moves the border to the right to make the new slide big long enough to quickly read it;
  2. he moves the border to the left to make the slides only just large enough that he can remember what it said from that quick read, leaving the speaker's head at very nearly the maximum size; and
  3. he listens to the speaker's discussion about the slide, until a new slide comes up.

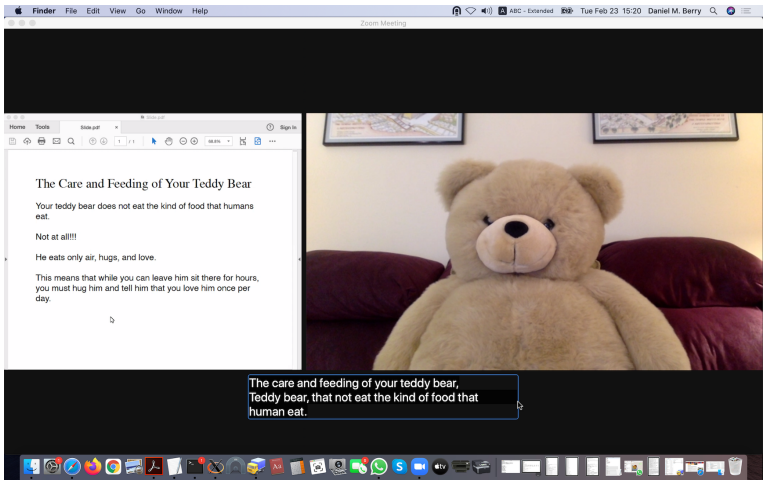

Fig. 2. AI-Generated Alive Captioning

In either case, if the deaf person finds a button labeled "CC Live Transcript" among the buttons at the bottom of his Zoom window, then he will usually find a caption window, already showing what any participant in the meeting has said. This caption window can be dragged to any place on the Zoom window to avoid its covering someone's lips or anything else that is important to see. See Figure 2[10]. If this caption window does not come up, then he needs to click on the "CC Live Transcript" button and select "Show Subtitle" in the menu that pops up. See Figure 3, which shows almost the same menu, but with "Hide Subtitle" where "Show Subtitle" would be if the captions were not already being displayed. If he does not want to see captions at all, then he needs to click on the "CC Live Transcript" button and select "Hide Subtitle" in the menu that pops up, as is shown in Figure 3. If he happens also to be recording the lecture, he can be assured that the caption window is not captured in the recording.

[10] The captions that you see in the figure are the result of my reading the title and the first item of the slide in the figure. It is clear that the AI-generated live captioning is not perfect. Part of the problem is that my speech is far from perfect, I do not pronounce correctly any sound that I cannot hear, and among the sounds that I do not hear is that of "s".

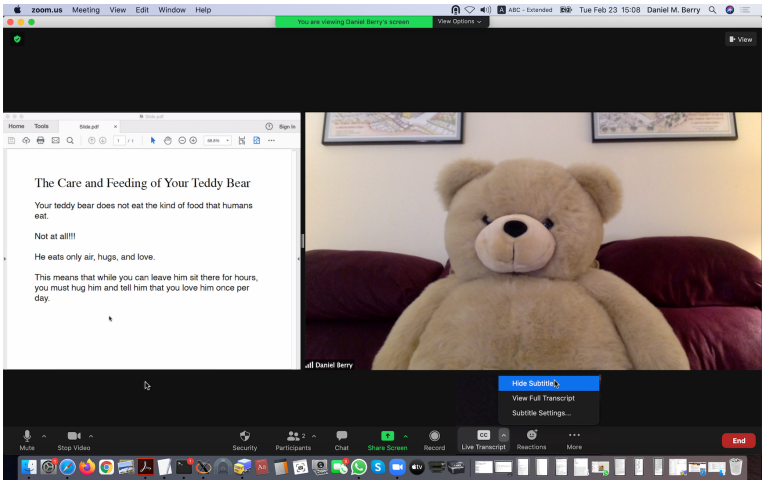

Fig. 3. Host Hiding Captions, a.k.a. Subtitles, After Enabling Alive Transcript

3 What Happens When the Slide Sharer Records a Zoom Meeting

This section describes what happens when the computer that is sharing slides in a Zoom meeting is the same computer that is recording the Zoom meeting. What happens explains why it is necessary that the sharing and the recording be done on two different computers in order that the resulting video be accessible to the lipreading deaf.

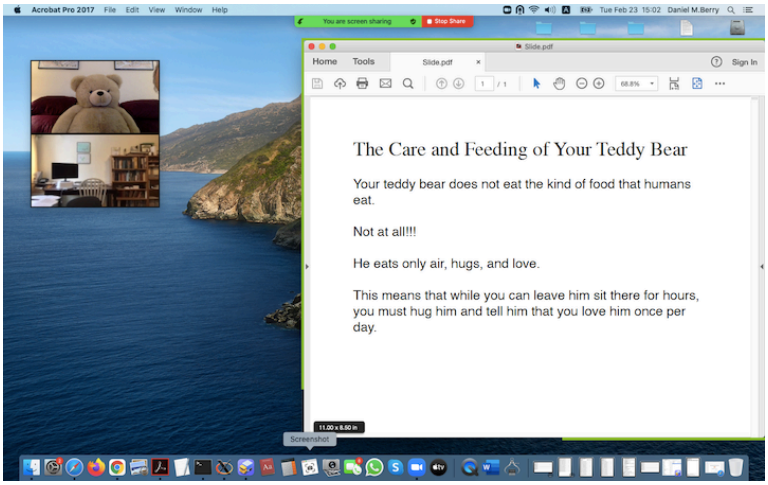

Fig. 4. Pieces of the Zoom Window Scattered Over the Monitor

In a Zoom meeting M, when any participant P, including the host, shares any artifact, the view on P's computer is different from those of all other participants. P does not see the standard divided Zoom screen, such as shown in Figures 1, 11, and 12. Instead, E sees the various pieces of the Zoom screen scattered all about the monitors available to em, as is shown in Figure 4. In addition, there is now a green border around the shared artifact unmoved from where it was from before the sharing. P interacts with the shared artifact according to the features of the artifact's application. There is really *no* Zoom window to speak of.

Suppose that this sharing P starts a local, not cloud, recording of M, either as the host or with permission of the host. Since there is no real Zoom window to capture in the recording, Zoom builds a simulated Zoom window that

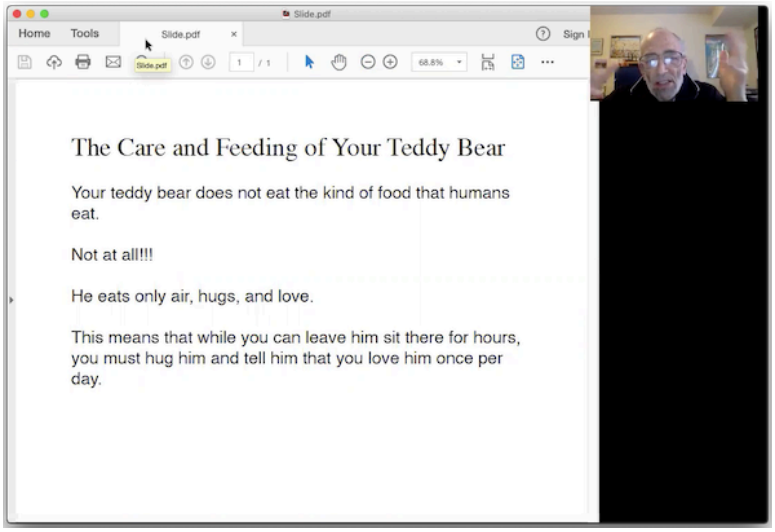

Fig. 5. The Simulated Zoom Window That is Built in the Recording

becomes the layout of the Zoom window that is immortalized into the recording. In any case, as shown in Figure 5, the simulated layout shows the shared artifact as big as possible on the left side of the Zoom window, with only a small sliver of space, SSS, on the right side of the Zoom window for faces. According to `https://support.zoom.us/hc/en-us/articles/360025561091-Recording-layouts`, what is recorded into the right side depends, in any instant during the recording, on P's current mode. If P is currently "hiding thumbnail video", then no thumbnail of any face is recorded into the right side of the Zoom window. If P is currently "showing active speaker", then a thumbnail of only the active speaker's face is recorded into the right side of the Zoom window. If P is currently "showing thumbnail video", then as many thumbnails as can be fit into the SSS are recorded into the SSS. The set of thumbnails recorded into the SSS probably[11] includes that of the speaker at the top of the SSS. Thus if any head is recorded, each head, including that of the speaker, is thumbnail sized, impossible to be lipread.

There appears to be no way to customize the layout of the simulated Zoom window when the recording is local, i.e., the video is generated on the recording computer and not in Zoom's cloud. The only way to avoid recording this default layout is for the recording to be done on a computer other than the one that is sharing the slides, which is usually the speaker's own computer so she can easily control the progression from slide to slide as she speaks. In this case, the recording captures the actual Zoom window on the recording computer. This observation is the basic requirement achieved by the instructions given below.

There may be some advantages to recording to the cloud, despite that cloud space is limited and recordings are saved for only 180 days. Zoom has recently added to its cloud recording the ability to record the active speaker, the gallery, and the shared screen separately. During the meeting, when the "Record" button is clicked, "Record to the Cloud" is selected instead "Record on this Computer". Later, after the recording is stopped, and the separate recordings are available, a video editing application can be used to composite a video with any desired layout, including those suggested in Figures 11 and 12, which would use only the active speaker and the shared screen recordings.

If in addition, the Zoom session being recorded is a webinar and not a meeting, then gallery mode shows the images of only the panelists who have unstopped their videos. If every panelist other than the speaker agrees to unstop his or her video only when he or she is asking a question, then gallery mode is a gallery of one when the speaker is speaking and is a gallery of two when a panelist is asking a question of the speaker. In this case, the compositor could use the gallery recording instead of the speaker recording. Doing so would allow seeing the speaker's reaction to a question. One limitation is that the more panelists there are who unstop their videos, the smaller are the lips of the person speaking in the gallery view.

[11] I have seen no Zoom, Inc. specification of this "probably" claim, and my experience is not quite enough to confirm the claim. However, I have never seen any case in which this claim is not true. On the other hand, I have never recorded such a lecture in which there were more face thumbnails than could fit into the SSS. So, I don't know what happens if the SSS overflows.

4 Making a Recorded Zoomed Lecture Accessible to the Deaf While Not Excluding the Blind

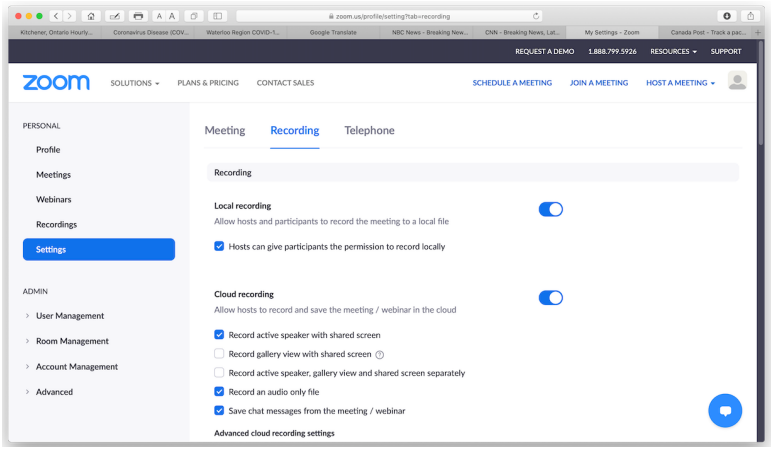

Fig. 6. Enabling Recording by Host and Participants

Prior to starting the Zoom meeting for the lecture to be recorded, the host logs into his Zoom account to set some parameters, which cover all future Zoom meetings — at least until the parameters are changed again. On the left size of the Zoom window, underneath "PERSONAL", he clicks "Settings", and then clicks on the "Recording" tab. In this tab, under "Local recording", he slides "Allow hosts and participants to record the meeting to a local file" on, and he checks "Hosts can give participants the permission to record locally"[12]. See Figure 6 for the results of these settings.

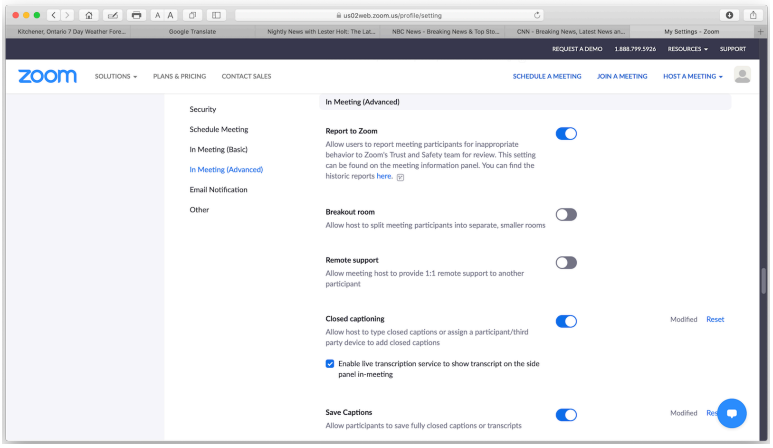

Fig. 7. Permitting Hosts to Enable Live Transcription That Can Be Seen by Participants

In addition, also within "Settings", he clicks on the "Meeting" tab. As is shown in Figure 7, in this tab, under "In Meeting (Advanced)", he slides "Closed captioning" on, and he checks "Enable live transcription service to show tran-

[12] In case a recording is ever to be saved to the Zoom, Inc. Cloud, in the same tab, under "Cloud recording", he slides "Allow hosts to record and save the meeting / webinar in the cloud" on, and he checks "Record active speaker with shared screen".

script on the side panel in-meeting". Doing so allows the host to enable, for any meeting, AI-generated live captioning that can be seen by any participant, as is described at the end of Section 2.

In the rest of this section, "muting" means "clicking on the 'Mute' button in the bottom left of the Zoom window to leave in its place an 'Unmute' button *with* a red line across it"; "unmuting" means "clicking on the 'Unmute' button in the bottom left of the Zoom window to leave in its place a 'Mute' button *without* a red line across it"; "stopping the video" means "clicking on the 'Stop Video" button in the bottom left of the Zoom window to leave in its place a 'Start Video' button *with* a red line across it"; and "unstopping the video" means "clicking on the 'Start Video' button in the bottom left of the Zoom window to leave in its place a 'Stop Video' button *without* a red line across it".

Now begin the per-meeting instructions.

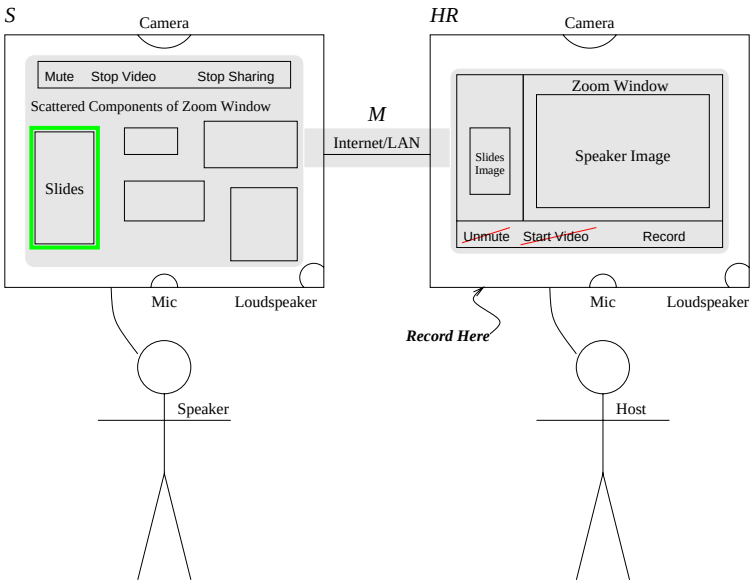

Fig. 8. Configurations of Computers Involved in Recording

As is shown in Figure 8, the speaker, who shares the slides, is on one computer, S, the host, who does the recording locally, is on another computer, *HR*. This host is the Zoom meeting host, who has a role different from that of the moderator, who might be on yet another computer. Because each role has enough to do during a recording, it is convenient that the speaker, the host, and the moderator be different persons. However, occasionally, the host and moderator are the same person. Also, if a speaker must prepare a video of her talk for later presentation, all three roles may be occupied by the same person, namely, the speaker. The next section discusses modifications to these instructions to deal with this special case.

1. Each of *HR* and S is connected to the Internet, and each is preferably connected via its modem, router, repeater, or hub by cable to minimize network delay to and from it.
2. On the *HR* Zoom window, the host starts a Zoom meeting, M, as the host of M. If M is a webinar, then the speaker, the moderator, and any deaf person among the participants need to be promoted to be in the panel so that they have all the features they need. To promote any participant named P, the host needs to click on the "Participants" button at the bottom of the *HR* Zoom window; in the resulting Participants window, hover over the name P until "More" shows up; click on "More"; and then click on "Promote to Panelist".
3. On the *HR* Zoom window, the host enables Auto-Transcript to make AI-generated live captioning visible to all participants in M. As is shown in Figure 9, he clicks on the "CC Live Transcript" button at the bottom of his Zoom window and then clicks on the "Enable Auto-Transcription" item in the menu that pops up. A caption window will come up, similar to the one shown in Figure 2, announcing, "You have turned on live transcription". Each other participant will now see a "CC Live Transcript" button at the bottom of er Zoom window and a caption window,

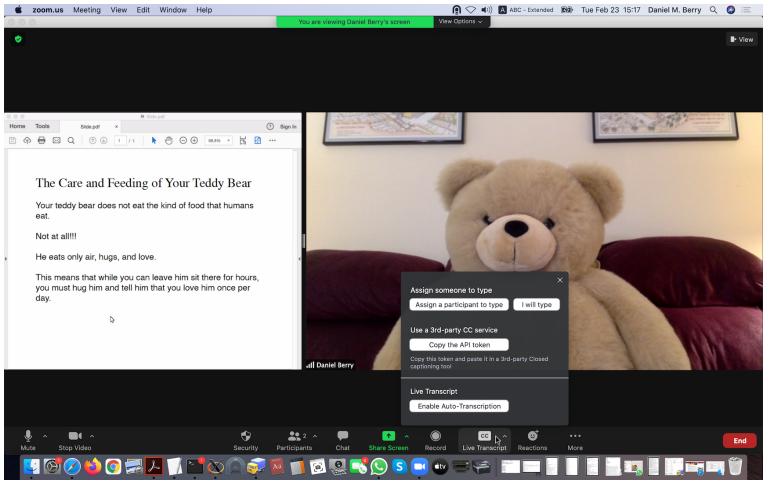

Fig. 9. Enabling Auto-Transcript to Get AI-generated Alive Captioning

showing a transcript of what is being said by any participant in M. If the host does not want to see the captions, he needs to click on the "^" in the button labeled "CC Live Transcript" and select "Hide Subtitle" in the menu that pops up, as is shown in Figure 3.

The host then announces to the participants:

(a) AI-generated live captioning has been turned on.

(b) The caption window that comes up can be moved to any place on the Zoom window to avoid its covering someone's lips or anything else that is important to see.

(c) If a caption window has not come up, then you need to click on the "CC Live Transcript" button at the bottom of your Zoom window and select "Show Subtitle" in the menu that pops up.

(d) If a caption window has come up, but you do not want it, then you need to click on the "CC Live Transcript" button at the bottom of your Zoom window and select "Hide Subtitle" in the menu that pops up.

(e) If you happen also to be recording the lecture, please note that the caption window is not captured in the recording.

Figure 2 shows the caption window. Figure 3 shows the menu that comes up when the "CC Live Transcript" button is clicked after Auto-Transcript has been enabled and the caption window is present. The top-most menu item shows "Hide Subtitle". If, however, the caption window is not present, then the top-most menu item shows "Show Subtitle" in the place of the "Hide Subtitle" item.

4. On the *HR* Zoom window, the host allows the speaker to share her slides by clicking on "Security", and then under "Allow participants to:" clicking on "Share Screen". See Figure 10.
5. On the S Zoom window, the speaker shares her slides, by clicking on "Share Screen" in the bottom center of the Zoom window and then selecting the artifact that she wishes to share from among the choices offered to her. She and the moderator, if E is not the host, then make small talk — maybe even tell some jokes to warm up the live audience ☺ — while the host finishes setting up the recording.
6. On the *HR* Zoom window, the host selects "Speaker Mode". At this point, the Zoom window on *HR* shows the speaker's slides on the left and the speaker's head on the right, as is shown in Figure 1.

   If the speaker has already delivered[13] an electronic copy of the slides that can be made available with the video,

   then, the host moves the border to the left to minimize the slides to a thumbnail no larger than necessary for a viewer be able to know where he should be in a local copy of the speaker's slides. The speaker's head is now as large as possible. This *HR* Zoom window layout, shown in Figure 11, is what will be immortalized into the recording.

[13] I have learned the hard way not to trust any speaker's promise to deliver the slides later. If you don't have the slide file in hand, act as if you will never have them. If you assume that you will have them, but don't eventually get them, the image of the slides in the recording will be too small for recording's viewers to read the slides.

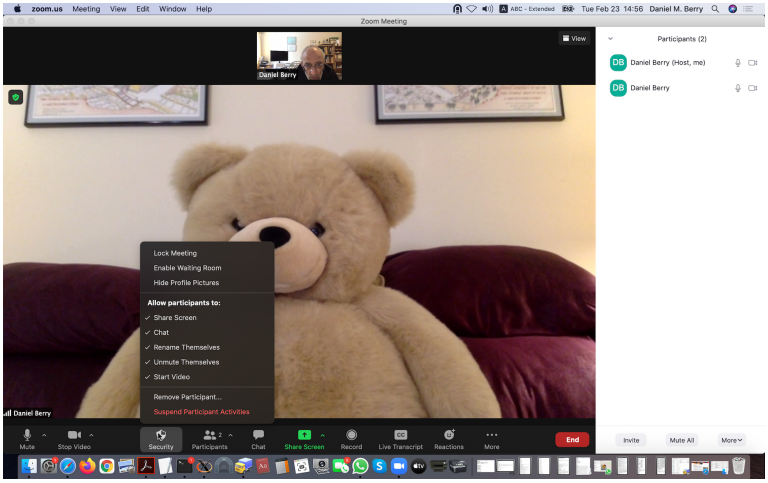

Fig. 10. Allowing Other Participants to Share Their Screens

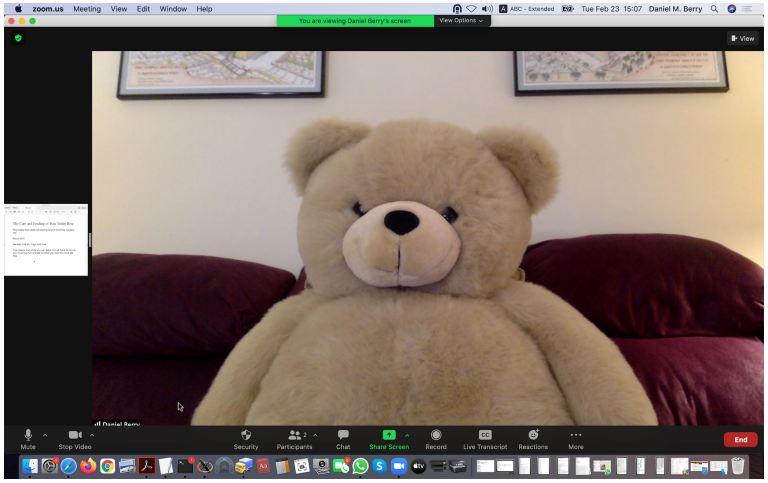

Fig. 11. Division of Zoom Window After Adjusting to Minimize Slides

Otherwise, the host moves the border to the left to minimize the slides to the smallest size at which all the slides are readable; the host should ask the speaker to move to the most detailed slides she has, so that he can calibrate the minimum size of the slide image. Usually the border will be near the middle of the Zoom window, making the slides and the speaker's head about the same size. This compromise is required to be sure that everyone who views the video will both (1) be able to read the slides and (2) have a chance to read the speaker's lips. This *HR* Zoom window layout, shown in Figure 12, is what will be immortalized into the recording.

Note that any change of the *HR* Zoom Window during the recording will be captured by the recording. So, the host should avoid moving the border during the recording. Nevertheless, if a slide image is too small to be read for its purpose, the border should be moved temporarily to allow the slide to be read for this purpose, and then moved back to where it was.

7. On the *HR* Zoom window, which is now in Speaker Mode, the host needs to do a few commands to try to prevent the active speaker window to the right of the border in his Zoom window, which is being immortalized into the recorded video, from showing any participant other than the speaker, except when a participant is speaking purposefully, e.g., asking a question of the speaker. The two events that cause Zoom to change the active speaker

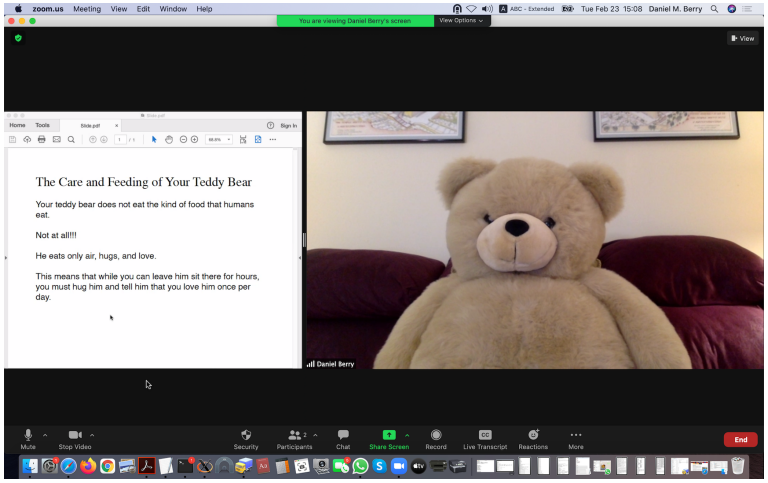

Fig. 12. Division of Zoom Window After Adjusting to Balance Slides and Speaker's Head

to another participant are (1) a participant's making a noise, even a sneeze, and (2) a new participant's joining M. The steps below are intended to allow these events to occur without Zoom's changing the active speaker.

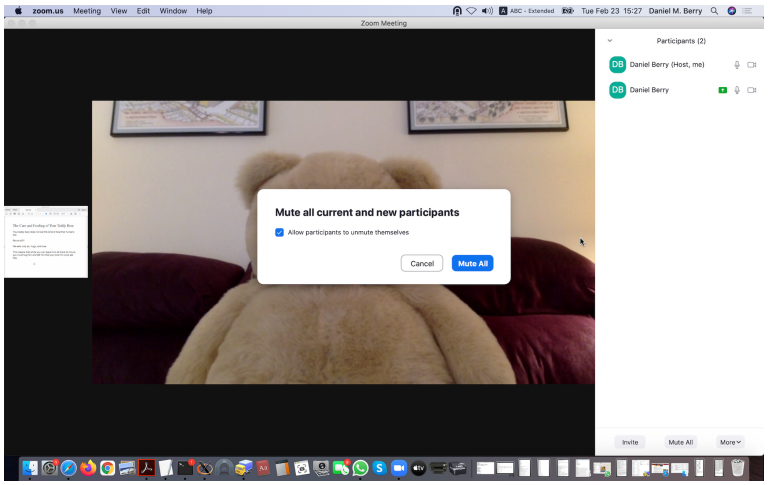

Fig. 13. Muting All Participants

(a) The host must mute himself and every participant but the speaker: The host clicks on the "Participants" button on the bottom of his Zoom window. The Participants list pops up in the right size of the Zoom window. At the bottom of this list, in the middle is a "Mute All" button. The host clicks on it and confirms "Mute All" in the popup window that appears, while leaving "Allow participants to unmute themselves" checked. See Figure 13. Suddenly, the mic icon for each participant, including the speaker but not including the host, in the Participants list gets crossed off. See Figure 14.

(b) The host asks the speaker to unmute herself by moving his mouse to the speaker's line in the Participants list and clicking on the blue "Ask to Unmute" that becomes visible. See Figure 15. As shown in Figure 16, the speaker sees that she has been asked to unmute herself. She does so. Alternatively, the host could just verbally ask the speaker to unmute herself.

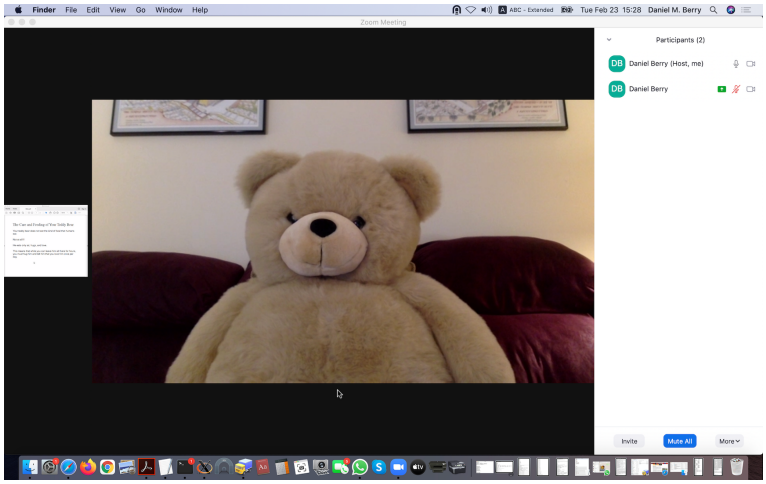

Fig. 14. All, Including the Speaker, Are Muted Except the Host

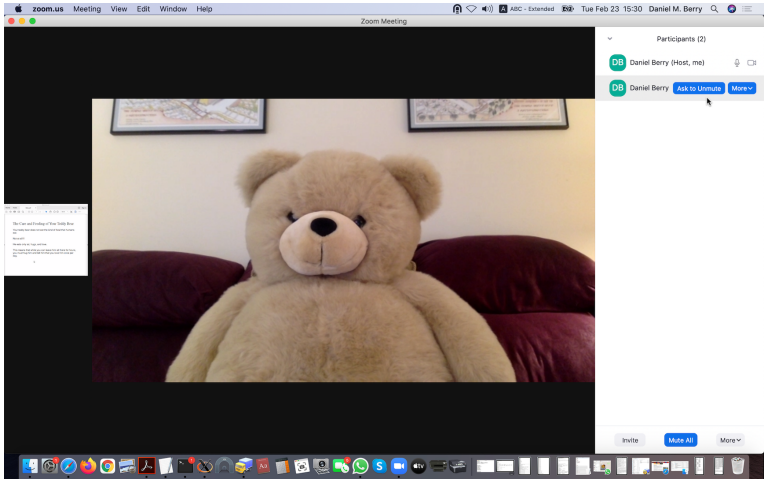

Fig. 15. Host Asking Speaker to Unmute Herself

(c) The host verbally tells all participants that they are now muted to help keep everyone's Zoom window focused on the speaker. However, if anyone wants to ask the speaker a question, E can unmute emself by clicking on the "Unmute" button at the lower left of er Zoom window or by holding down the space bar while speaking. When the questioning is finished, E should mute emself by clicking on the "Mute" button that used to be the "Unmute" button. If E forgets to mute emself, the host can mute em at er entry in the Participants list.

(d) The host verbally requests all participants to please also unstop their videos on their Zoom windows whenever they unmute themselves to ask questions or make comments. Doing so will allow their speaking to be lipread, as they become the active speaker. There will be participants who cannot comply with this request, because they have connected to M by a device that has no usable camera. The moderator should repeat with er face visible any question asked by participant who does not show er face.

(e) With all the verbal instructions done, the host mutes himself. As a result, as shown in Figure 17, all participants, including the host, but not including the speaker are muted.

(f) The host clicks on "Participants" button to get rid of the participants list in the right hand side of the *HR* Zoom window.

(g) The host switches his Zoom window to Gallery Mode.

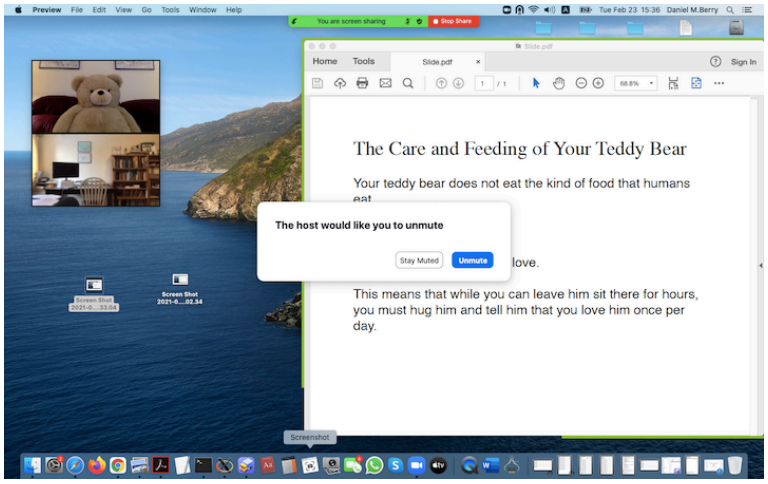

Fig. 16. Speaker Sees That She Has Been Asked to Unmute Herself

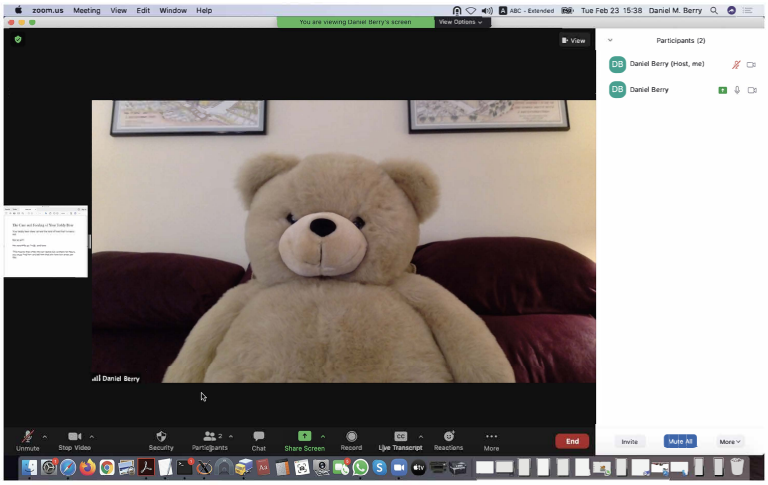

Fig. 17. All, Including the Host, Are Muted Except the Speaker

(h) In Gallery Mode, the host hovers his mouse over the upper right hand corner of the speaker's thumbnail until the three-dots button appears. The host clicks on that button and selects and clicks on the "Pin" menu item. See Figure 18. The result is that the *HR* Zoom window is back into Speaker Mode and the pinned window stays the active speaker until it is unpinned. The pinned window stays the active speaker even when another participant speaks or a new participant joins M.

(i) The host leaves the speaker pinned until the first time someone wants to ask a question. At this time, the host unpins the speaker by clicking on the "Remove Pin" button that becomes visible when the host's mouse hovers over the upper left corner of the active speaker's image. See Figure 19. The unpinning leaves the host's Zoom window in Speaker Mode, under control of Zoom's speaker changing algorithm. By this time, the frequency of new participants joining M should have decreased to zero, and Zoom will properly display whoever is currently speaking in the active speaker window. Most importantly, the recorded video will be capturing the active speaker at all times.

8. On the *HR* Zoom window, the host clicks on the "Record" button on the bottom middle of the Zoom window, choosing to "Record on this Computer". See Figure 20.
9. On the Zoom window in some computer that is attending M, the moderator prepares to introduce the speaker. E unstops the video to allow er own face to show. Also, E will unmute or mute according to when he needs

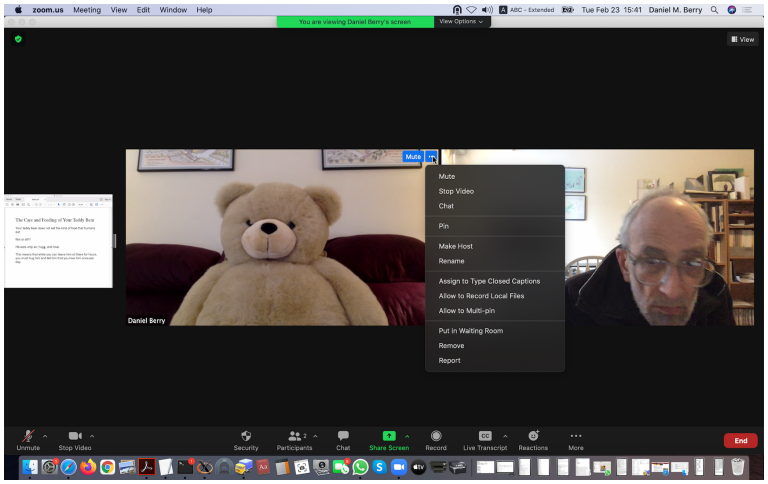

Fig. 18. Pinning the Speaker

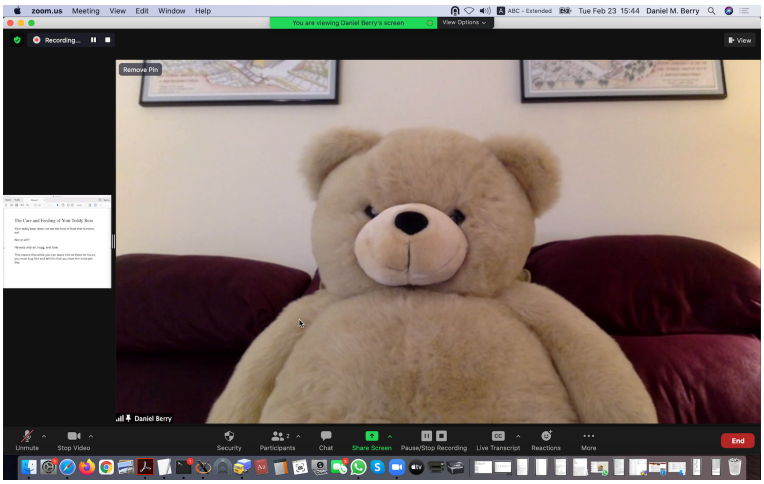

Fig. 19. Unpinning the Speaker

to be actively moderating the speaker or just listening, respectively. The moderator starts the actual lecture by introducing the speaker and then letting her take it away.

During the lecture, moderator watches for signs that a participant wants to ask a question. When a participant wants to ask a question, the moderator unmutes emself and unstops er video. E asks the questioner to unmute emself and to unstop er video and then to ask the question. When the give and take over the question is done, the moderator asks the questioner to mute emself.

In general, the moderator should try to slow down the speed of the switching between people speaking up to allow Zoom to catch up and be displaying in the active speaker window the person who is actually speaking, and so that everything important that anyone says is said while er face is being shown in the speaker view.

10. On the S Zoom window, the speaker gives her lecture while manipulating the shared slides. Note that the slide show application window should not be moved out its current monitor without stopping the sharing, moving the window, and then starting up a sharing of the moved window. Moving a shared window to another monitor results in freezing the image of the shared document, as seen by all participants but the speaker and as recorded, to the image that was showing at the time of the move.
11. When the lecture and the Q & A are completed, on the *HR* Zoom window, the host stops the recording, by clicking on the black or white — depending on the background color — square button to the right of the flashing circled

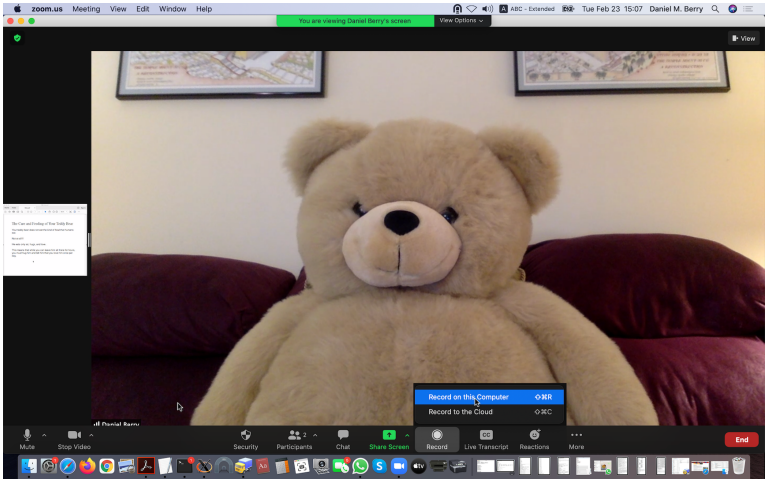

Fig. 20. Starting the Recording on *HR*

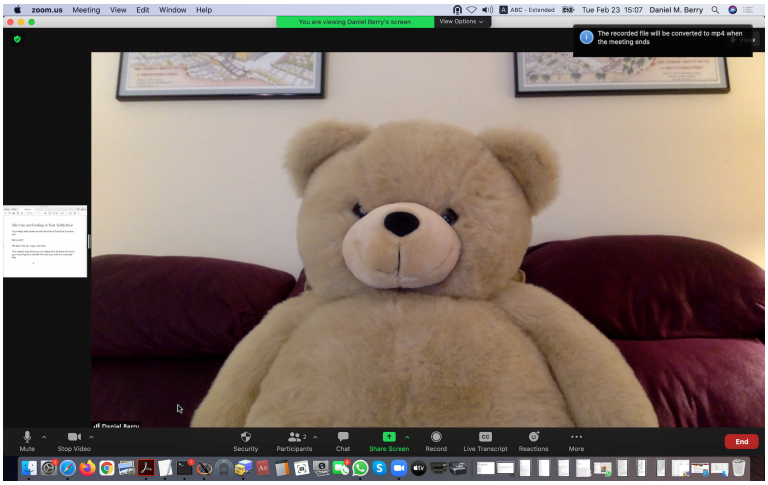

Fig. 21. After Recording Has Been Stopped

red dot and "Recording . . . " indicator in the upper left of the Zoom window. See the upper left corner of Figure 19 (which actually comes *after* Figure 20 in the history of M). The host takes note of the information that "The recorded file will be converted to mp4 when the meeting ends". See Figure 21.

12. The moderator thanks the speaker and says "Good-bye." to all.
13. On the *HR* Zoom window, the host clicks on the red "End" button in the lower right of the Zoom window and confirms to "End Meeting for All". See Figure 22. He does not click on "Stop Converting" when a pop up window informs him that Zoom is "Converting meeting recording" and "You have a recording that needs to be converted before viewing". Note that the conversion can consume several minutes. See Figure 23, showing both the beginning of the conversion and near the end, several minutes later.
    When the conversion is done, another pop up window asks whether to "Save my recording at" some file in a directory containing a lot of Zoom generated files. The host clicks "Save" to save the .mp4 file. The directory with the saved .mp4 file pops up allowing the host to play the .mp4 file. Figure 24 shows both the request to save and the resulting directory.

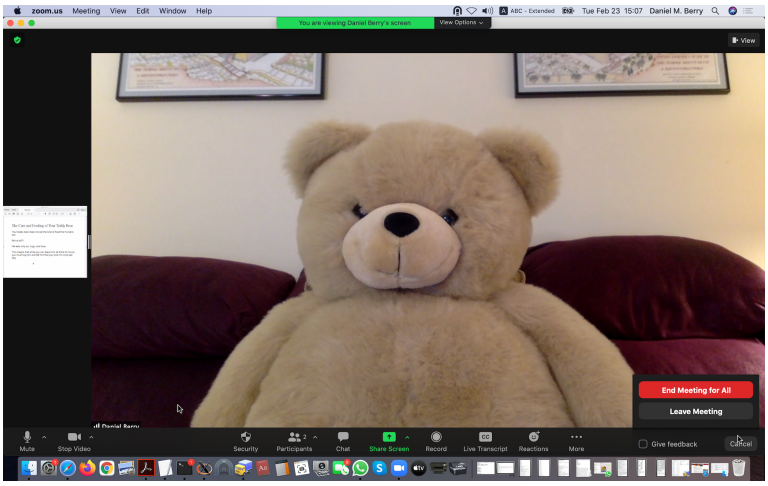

Fig. 22. Ending the Meeting

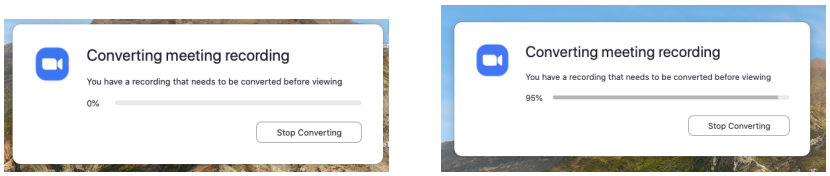

Fig. 23. Converting the Recording to an MP4 Video: From 0% Through 95% and Beyond Over Many Minutes

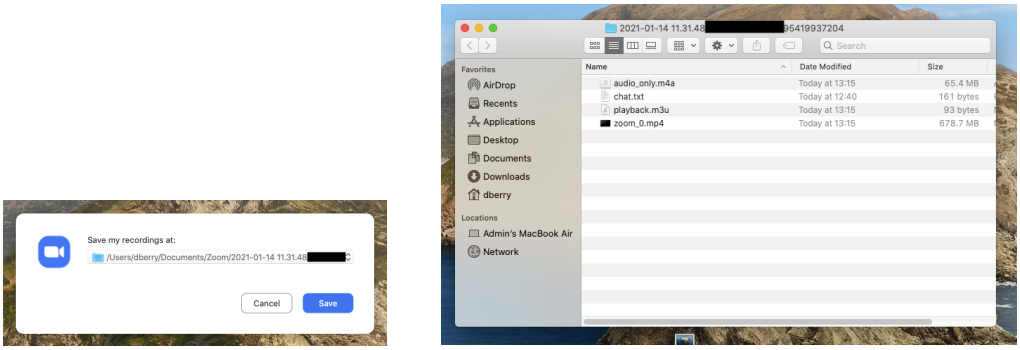

Fig. 24. Saving the Recording

5 Other Scenarios

Sometimes a speaker, scheduled to speak, e.g., at a conference, has to record her own lecture to produce a backup video. The intent is that this backup video can be run by the conference organizers in place of her giving her lecture alive at the conference should she not able to connect at the scheduled time of her lecture or if the connection is so poor that her lecture will be difficult to follow.

The speaker will need two computers S and *HR* and she will play also the role of host. If at all possible, the two computers should be connected by wire to the same high-speed LAN. There will be no one playing the role of moderator, since the conference organizers will introduce the speaker before playing the pre-recorded video.

The two computers should be in two different rooms, such that the mic of one machine does not hear the sound emitted by the loudspeaker of the other machine or the sound of a human being speaking at the other computer. It is important to ensure that the only sound signal that is captured in the recording is that coming through the channel to the sound output of the recording computer. It is still helpful if the Zoom window on the recording computer is muted and the video is stopped.

The speaker follows the procedure of Section 4 with the speaker playing the host role and with the following modifications:

2 There is no need for the webinar steps, since a recording session is not a webinar.
3 There is no need to enable Auto-Transcript, since does not need them to know what she is saying, and in any case, the captions do not get captured into the recording.
5 There is no need for the speaker and the non-existent moderator to make small talk or tell jokes.
6 How the Zoom window on *HR* is divided should be decided according to whether the speaker's slides will be available before the conference for download into attendees' computers. In the usual case, they probably will not be available, or many listeners will not have downloaded ahead of time. So it's probably best to take the "Otherwise" choice, setting the border to be near the middle of the Zoom window.
7 This step and all of its substeps are skipped, as there are no other participants to cause the active speaker to change.
9 This step is skipped, as there is no moderator.
12 This step is skipped, as there is no moderator, and there is no reason to record a "Good-bye" into the video.

The Appendix gives the actual list of instructions modified as described above. This list can be used directly with no need for on-the-fly interpretation.

After the recording is saved as a video, the speaker might wish to edit the video to remove the dead time at the beginning and the end of the video.

6 Instructions to Send to Speaker in Advance of the Lecture

Below is the skeleton of a letter that can be sent a speaker in advance of her lecture to try to get her to give her lecture with the highest bandwidth connection that is possible given her network connection and to provide her slides in advance of the lecture. Adapt it to your specific circumstance.

> Dear Speaker;
>
> Thank you for agreeing to speak at XXXXX on YYYYYY on 2 Nunember 2021 in Somewheresville.
> This letter discusses some matters that concern helping all attending your lecture be able to see and hear it.
> Some members of your audience will be lipreading. As much as possible, please keep your camera on, and face the camera, so that your audience can see your eyes and mouth. To help improve the refresh rate of your image to make lipreading easier, if at all possible, please use a wired connection or sit close to the router. We do understand that your specific work environment may not permit a wired connection or being close to the router.
> The lipreaders will be maximizing the image of your face at the cost of minimizing the image of your slides. If it is OK with you, please send to us a copy of your slides so that they can follow your talk with a larger local copy of your slides. A local copy of your slides also allows the blind to follow your slides with their reading software. If you prefer, we can limit the distribution of these slides to the lipreaders and the blind.
>
> Thank you

7 Closed Captioning of the Video

Once the video is completed, it should be closed captioned. Beware, because the quality of closed captioning (CC) varies according to how it is done. The best CC is that inserted during post production of the video by a skilled human transcriber, who listens to the speaker on the video and produces captions with correct spelling and punctuation, and maybe even with nice formatting. This transcriber should probably be someone who knows the topic of the lecture so that he or she knows the vocabulary, the acronyms, how to pronounce any formulae, etc. Otherwise, technical terms will be misunderstood and transcribed — often very funnily — as something totally unrelated to what was actually said. The desired quality of CC is very labor intensive and expensive.

The more usual substitute for manual CC is automatic CC (ACC), such as can be provided by Google for a video uploaded to YouTube. In ACC, an artificial intelligence listens to the sound coming through its audio channel and applies the results of machine learning to produce a transcript of varying quality. The main drawback of ACC is that, being driven by AI, ACC is significantly less accurate than lipreading or real-time manual closed captioning, which is driven by the RI (real intelligence) that resides in human beings. Humans know how to make use of the semantic context that an AI is oblivious to. When the speaker speaks very clearly, such as do TV news anchors, ACC is quite good and it helps a lipreader get through periods during which low network bandwidth causes the video to be degraded in favor of preserving the sound quality. In this situation, lipreading becomes impossible. ACC works because the sound quality has been preserved. However, if the speaker speaks with a non-native English speaker's accent, ACC can be quite poor, to the extent that a deaf listener will understand more from being able to read the speaker's slides than from reading the garbled captions that bear little relation to what the speaker says. So it will often be necessary to download the results of ACC on a video, have a human edit it manually, and then upload the edited captions into the video to replace the ACC. Again, this human being should be someone who knows the topic of the lecture, particularly if the lecture involves formulae. Otherwise, the resulting captions tend to be only marginally better than those produced by ACC with a native English speaker, and formulae will be totally mangled.

Even at its best, closed captioning (CC) should not be considered a complete replacement for allowing the deaf to read lips. Just as a hearing person can hear the tone and emotions of a voice, a skilled lipreader can read the tone and emotions of the lips and face being lipread. Thus, no more than a hearing person would be satisfied with just a transcript of a lecture, would a lipreading deaf person be satisfied with just the captions. In addition, when a speaker says the pronoun "this", while moving the mouse to point to a specific item on the shared slides, there are visual clues in the speaker's face and shoulder movement that direct the listener's eyes to look at where the pointer is in the image of the shared slides that identifies the referent of "this". This information is lost if only the captions of what the speaker says are available.

The three basic steps of a partially automated way to closed captioning an already recorded video, V.mp4, are

1. Submit V.mp4 to some AI-assisted captioning service to obtain a first-draft captioning in the form of an .srt file, V.srt.
2. Edit the V.srt, which happens to be a plain text file, to correct all the mistakes that the AI-assisted captioner made.
3. Either
   (a) make both V.mp4 and V.srt available to be played by an .mp4 player that can figure out, or can be told how to link the two files together, as the video is being played, or
   (b) submit V.mp4 and V.srt to a program that creates a new video Vcaptioned.mp4, with the captions burned into the new video.

There are a number of ways to achieve Step 1 of automatically obtaining a first-draft captioning: Submit the video V.mp4 to any of

- `https://www.amberscript.com/`,
- `https://otter.ai/`,
- `https://speechtext.ai`,
- `https://streamer.center/` (not yet, the company says, but in about a month),
- `https://studio.youtube.com/`,
- `https://zubtitle.com/`, and

others. For each such service, you need to have an account.

The service will generate a close caption file that you can export as an .srt file, *V*.srt. This file is plain text; so it can be easily edited with an ordinary text editor. Even if the transcription is poor, it is worth obtaining the resulting .srt file, because it attaches a serial number and beginning and ending timestamps to each phrase of the captions, e.g.,

```
12
00:01:57,000 --> 00:02:01,760
and manipulates the interface
phenomena as specified, will
```

says that "and manipulates the interface phenomena as specified, will" is the 12th phrase, and it is displayed during the time span from 0 hours, 1 minute, 57 seconds, 0 milliseconds to 0 hours, 2 minutes, 1 second, 760 milliseconds.

Step 2 can be achieved simply by editing the *V*.srt with your favorite text editor. Just do not change any serial number and any timestamp. Neither add nor remove any blank line. It's hard to imagine any correction of the text, reflecting what is actually said in the video, that results in text that is too long for its time duration.

If Option a of Step 3 is taken, then the viewer will need to use an application, such as VLC, whose commands allow specifying any .srt file as the caption file for the .mp4 file that is about to be played.

Option b of Step 3 can be achieved by submitting *V*.mp4 and *V*.srt to an application, such as Adobe Premiere Pro. Unless the bitrate of the output is chosen carefully, the resulting *V*captioned.mp4 can easily be an order of magnitude bigger than *V*.mp4. You must use the "Match Source – Adaptive Medium Bitrate" setting. With this setting, the size of the resulting *V*captioned.mp4 is a bit bigger than the sum of the sizes of *V*.mp4 and *V*.srt.

The files for the recordings of two fun lectures by Daniel Berry at the online appendix show the files that result from both options of Step 3.

Nick Anisimov describes the whole process at `https://nck-anisimov.medium.com/how-to-create-an-srt-file-the-complete-guide-6033f85dad29`.

There are even integrated solutions that assist in the whole process, e.g., `https://www.nchsoftware.com/videopad/`.

For basic instructions on how to turn on ACC on videos uploaded to Youtube, see

`https://support.google.com/youtube/answer/6373554?hl=en`.

For instructions on downloading the results of ACC on a video uploaded to Youtube, see

`https://titaniumhelp.fullerton.edu/m/100674/l/498105-how-do-i-export-the-youtube-auto-generated-captions`

For instructions on how to upload manually edited captions to a video uploaded to Youtube, see

`https://support.google.com/youtube/answer/2734796?hl=en`

Finally, keep in mind that, unlike while attending a live lecture, a person who has difficulty understanding the speaker at any point while listening to a video can pause the video, back up, and listen again. Thus, 100% perfection in captioning is not required.

Acknowledgments

I thank Maura Grossman for being willing to participate as a remote speaker in several of my tests of whether what I wrote here actually works. I thank Paul Eggert, Ian Goldberg, Ilse Mozga, Naomi Nishimura, and Joe Petrik for their comments on earlier drafts of these instructions. I thank Jared Cubilla for doing the manually edited closed captioning for the example recorded lectures that are at the Web site for the supplemental materials for these instructions and for answering my questions about the process. I thank Christopher Batty, Mike Godfrey, and Karthik Velakur for their help concerning adding captions to videos and compressing the resulting *HUGE* videos to a manageable size. Finally, I thank Martin Glinz for his discovery and documenting of the two-computer method described herein to make hearing-impaired accessible recordings of Zoom lectures. So cool!!!

A Appendix: Self-Recording Instructions

Sometimes a speaker has to record her own lecture to produce a video.

The speaker will need two computers S and *HR* and she will play also the role of host. If at all possible, the two computers should be connected by wire to the same high-speed LAN. There will be no one playing the role of moderator, since the conference organizers will introduce the speaker before playing the pre-recorded video.

The two computers should be in two different rooms, such that the mic of one machine does not hear the sound emitted by the loudspeaker of the other machine or the sound of a human being speaking at the other computer. It is important to ensure that the only sound signal that is captured in the recording is that coming through the channel to the sound output of the recording computer. It is still helpful if the Zoom window on the recording computer is muted and the video is stopped.

1. Each of *HR* and S is connected to the Internet, and each is preferably connected via its modem, router, repeater, or hub by cable to minimize network delay to and from it.
2. On the *HR* Zoom window, the speaker starts a Zoom meeting, M, as the host of M.
3. **This is a place holder in order that the subsequent steps have the same step numbers as in the full procedure.**
4. On the *HR* Zoom window, the speaker allows herself on the S Zoom window to share her slides by clicking on "Security", and then under "Allow participants to:" clicking on "Share Screen". See Figure 10.
5. On the S Zoom window, the speaker shares her slides, by clicking on "Share Screen" in the bottom center of the Zoom window and then selecting the artifact that she wishes to share from among the choices offered to her.
6. On the *HR* Zoom window, the speaker selects "Speaker Mode". At this point, the Zoom window on *HR* shows the speaker's slides on the left and the speaker's head on the right, as is shown in Figure 1.
   If the speaker *knows for sure* that an electronic copy of the slides will be made available the eventual viewers of the video,
   then, the speaker moves the border to the left to minimize the slides to a thumbnail no larger than necessary for a viewer be able to know where he should be in a local copy of the speaker's slides. The speaker's head is now as large as possible. This *HR* Zoom window layout, shown in Figure 11, is what will be immortalized into the recording.
   Otherwise, the speaker moves the border to the left to minimize the slides to the smallest size at which all the slides are readable; the speaker should move to the most detailed slides she has, so that she can calibrate the minimum size of the slide image. Usually the border will be near the middle of the Zoom window, making the slides and the speaker's head about the same size. This compromise is required to be sure that everyone who views the video will both (1) be able to read the slides and (2) have a chance to read the speaker's lips. This *HR* Zoom window layout, shown in Figure 12, is what will be immortalized into the recording.
7. **This is a place holder in order that the subsequent steps have the same step numbers as in the full procedure.**
8. On the *HR* Zoom window, the speaker clicks on the "Record" button on the bottom middle of the Zoom window, choosing to "Record on this Computer". See Figure 20.
9. **This is a place holder in order that the subsequent steps have the same step numbers as in the full procedure.**
10. The speaker quickly moves from the room where *HR* is located to the room where S is located.
    On the S Zoom window, the speaker gives her lecture while manipulating the shared slides. Note that the slide show application window should not be moved out its current monitor without stopping the sharing, moving the window, and then starting up a sharing of the moved window. Moving a shared window to another monitor results in freezing the image of the shared document, as seen by all participants but the speaker and as recorded, to the image that was showing at the time of the move.
11. When the lecture is completed, the speaker quickly moves from the room where S is located to the room where *HR* is located.
    On the *HR* Zoom window, the speaker stops the recording, by clicking on the black or white — depending on the background color — square button to the right of the flashing circled red dot and "Recording ..." indicator in the upper left of the Zoom window. See the upper left corner of Figure 19 (which actually comes *after* Figure 20 in the history of M). The speaker takes note of the information that "The recorded file will be converted to mp4 when the meeting ends". See Figure 21.
12. **This is a place holder in order that the subsequent steps have the same step numbers as in the full procedure.**

13. On the *HR* Zoom window, the speaker clicks on the red "End" button in the lower right of the Zoom window and confirms to "End Meeting for All". See Figure 22. She does not click on "Stop Converting" when a pop up window informs her that Zoom is "Converting meeting recording" and "You have a recording that needs to be converted before viewing". Note that the conversion can consume several minutes. See Figure 23, showing both the beginning of the conversion and near the end, several minutes later.
    When the conversion is done, another pop up window asks whether to "Save my recording at" some file in a directory containing a lot of Zoom generated files. The speaker clicks "Save" to save the .mp4 file. The directory with the saved .mp4 file pops up allowing the speaker to play the .mp4 file. Figure 24 shows both the request to save and the resulting directory.