

The Power of the Voice in Facilitating and Maintaining Online Presence in the Era of Zoom and Teams

**** On the Internet ****

August 2023 – Volume 27, Number 2
<https://doi.org/10.55593/ej.27106int>

Michael Cribb

Coventry University, UK
<aa6177@coventry.ac.uk>

Abstract

With the lockdowns of the COVID-19 pandemic and increasing popularity of video-conferencing software such as Zoom, the move to online and /or hybrid teaching has never been more rapid. With this change, however, maintaining presence in the classroom has become a great challenge simply because of the nature of online teaching. Presence is a teaching quality that enables the teacher to ‘own the room’ and create an atmosphere of focus and inspiration. With the loss of face-to-face contact and the diminution of body language that online teaching entails, the teacher has to rely more and more on their own voice to hold presence in the class. While voice has always been an important tool in the teacher’s expressive armoury, it takes on a more central role in online teaching and can be the only element that connects teachers to students. Yet many teachers still front classes where voice audio quality is severely restricted due in part to poor choice of microphone and setups on their behalf. In this article I will discuss the notion of presence in online classrooms with regard to voice, and show how teachers can maintain and manipulate this feature in order to retain appeal for students.

Keywords: presence, voice, online teaching, VARS, hybrid teaching

The recent COVID-19 pandemic and the resulting national lockdowns starting in 2020 meant that most English language classes, and in many cases whole courses, moved to online teaching. The vast majority of English language teachers needed to rapidly acquaint themselves, either willingly or unwillingly, with video-conferencing software (VCS) such as Zoom or Microsoft Teams in order to deploy these classes. Even though the lockdown restrictions have now been lifted in many countries, some teachers have continued with online teaching in one form or other, as is the case at my institute, either as stand-alone classes or

hybrid classes in which face-to-face (F2F) teaching in class is combined synchronously with online teaching.

This rapid shift to online teaching has brought with it a certain set of additional challenges over and above the traditional competencies that teachers need for F2F delivery. These have been documented as electronic classroom interaction competencies (e-CIC) through a number of studies (Moorhouse, Li and Walsh, 2021; Wells et al., 2022) and include technological, environmental and teacher interactional competencies. One such competency that spans all modes is maintaining a presence in the classroom. *Presence* is the ‘ability to command a room, hold pupils’ attention, and create an atmosphere of focus and inspiration’ (Hardy, 2021). It is often framed as ‘owning the room’ (Su & Wilkins, 2013) or ‘having gravitas’ (Newton, 2019).

Presence in online teaching is a complex amalgamation of human and situational factors but voice quality is a key component of this construct. The term *voice quality* in this article is taken to mean the clarity and intelligibility of a speaker over VCS as perceived by a listener (rather than the more narrowly-defined linguistic usage). This is less to do with the pronunciation and articulation of words by a speaker and more to do with the acoustic setup. The human voice has become much more important in creating and maintaining presence due to the teacher’s loss of body language and physical manifestation that inevitably occurs with online teaching.

The aim of this article is to demonstrate the significance of voice in creating and maintaining presence in the online classroom. After establishing the importance of voice quality in online teaching contexts, the article will introduce the Voice Audio Rating Scale (VARs), which is a means to measure and classify the audio quality of the teacher’s voice over VCS, and show how this scale can be employed as a learning tool with teacher-trainees. The article will also look at *hybrid* teaching environments (Nørgård, 2021) in which a teacher is simultaneously teaching to a cohort of students F2F and a cohort online. The juxtaposition of two sets of students in different teaching spaces means that maintaining presence with both sets entails a complex cognitive stance that teachers need to adopt in this context if they are to be successful.

Presence

Presence is a hard-to-define competency in teaching practice. It is sometimes described as the ability to ‘command a room and hold students’ attention’ (Hardy, 2021) or simply ‘being there’ (Lehman & Conceição, 2010). When a teacher has presence in the classroom, they have the ability to speak and interact with confidence, and the focus of the lesson is centred around them. Presence is perhaps not something which comes naturally to the human person, but most teachers understand the need for this early in their careers and develop a robust, but individual, classroom presence over their years of training and practice. Teacher-trainees are invariably given tips and guidance early in their careers on how to create presence through the use of body language, the projection of voice, the use of the classroom space, being passionate, etc. (e.g., Halonen, 2002).

But presence is not just simply a means to communicate with and control the classroom. Presence also has deeper roots in projecting sincerity and authenticity to the students. Umpleby (2014, p. 28) believes that it is from ‘an existential core, rather than through teaching performance, that teachers are able to connect with their students at a human level’. Classroom presence is a ‘mode of being’ (p. 28) in the fullest and deepest sense rather than a set of classroom competencies. In other words, presence is not just something that we *project*; it is an essential human quality that *connects* teachers to students. Teachers are ‘psychologically,

emotionally, and behaviorally present when they connect with others in an authentic way during the online learning experience' (Lehman & Conceição, 2010, p. 9). With the move to more and more online and distant digital platforms, the questions of how we can create and maintain this 'mode of being' has never been more pertinent. How can we maintain a 'spirit of care, authenticity, immediacy and empathy' with our students that they deserve, as Umpleby advises:

As man looks to technology for the answers, those human traits of 'being with another' infused with a spirit of care, authenticity, immediacy and empathy often seem to have little value and yet their transformative potential is priceless (Umpleby, 2014, p. 26).

We can visualise the dilemma for the teacher in Figure 1. On the left, we see a teacher in the traditional class, face-to-face and physically present with their students. The teacher has available to themselves the full range of body language, voice, and classroom space to engage and connect with the students. They can project authenticity, immediacy and empathy and be with the students in the fullest sense. On the right, however, we have the same teacher in an online VCS environment such as Zoom or Teams. The teacher's body language and facial expressions have been reduced to a 2-D representation in a small area in the top-right of the screen, which the online students may choose to ignore over the more imposing PowerPoint slides in the centre of the screen. Even when the students do look at the teacher, the delay in transmission usually means that gestures and expressions are not accurately timed with the voice being transmitted. The teacher has also lost their avenue for spatial movement: there is nowhere for them to move.

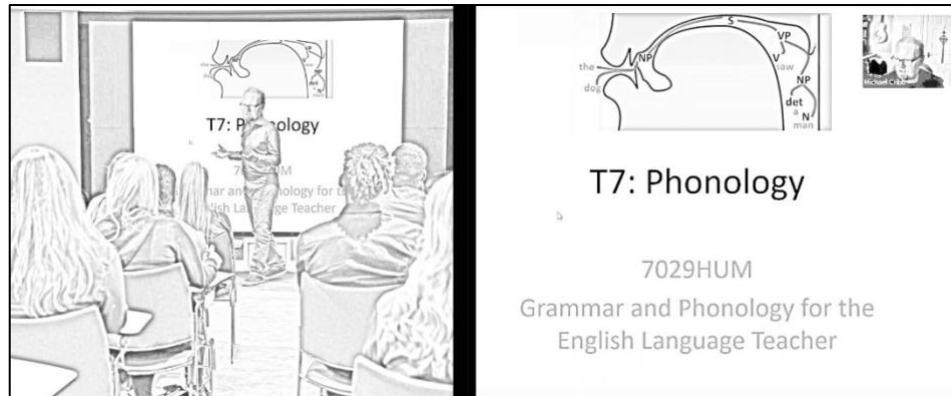


Figure 1. F2F v Online teaching

Due to this reduction in the visual projection and loss of space, the teacher's voice is effectively magnified and becomes the key component for holding presence in the online class. The voice, devoid of most gestures and facial expressions, becomes the dominant channel for communication. However, presence, as I have already outlined, is not just projection, but also connection. In the above diagram on the right, we note that the students' faces and voices do not show which means that the teacher cannot see how the students are reacting to what is being said. Subtle cues in F2F classes normally enable the teacher to pick up on any misunderstandings or confusions, reactions to asides and quips, and the level of interest or boredom. Without this feedback, the dreaded 'wall of silence' often descends over the class

where absolute silence is presented to the teacher forcing them to teach effectively, in NASA speak, ‘in the blind’.

Voice

In online teaching contexts, as the illustration above demonstrates, the human voice carries an extra burden in maintaining presence because of reduced visual cues, loss of space and disconnectedness with the students. It becomes one of the most powerful tools in the teacher’s expressive armoury for creating and maintaining presence in the classroom, if it wasn’t already so.

Voice is the sound produced at the mouth when we speak due to the action of the speech organs such as the tongue, larynx, etc. The human voice begins in the chest and diaphragm as air is expelled through the larynx and the vocal tract, i.e., oral and nasal cavities (Mahendru, 2014). The quality of the voice depends largely on the shape of the vocal tract and is modified by the physical arrangement of the pharynx, tongue, velum and lips among other parts. In normal speech we produce ten to fifteen speech sounds per second (Port, 2007) and the contrastive nature of these sounds enables the speaker to produce words and sentences that are intelligible to the listener. On ‘top’ of these sounds, we also employ prosodic features to package the speech stream in terms of pitch, rhythm and stress to the listener. Studies have shown how important this prosody is for the full intelligibility of speech and for the perceived liveliness of voice (e.g., Hincks, 2005). Most teachers develop a melodious voice over their teaching years, using sweeping intonation rises and falls, consistent rhythm and contrastive stress to maintain interest and liveliness for the students in what might otherwise be a long and tedious hour. We can also accompany the speech with paralinguistic features such as coughs, marked tones, etc. as well as varying the speed, tempo and loudness of the delivery.

In F2F teaching, the voice is an infinitely variable and immediately present tool with which the teacher can construct and maintain dialogue with their students. The voice not only carries most of the instructional content, but also enables subtle shifts in tones and emphasis to signal asides, praise, admonishments, jokes, banter, etc. Shifts in tone, emphasis and speed can foreground or background information and enable the teacher to signal the hierarchy of information (Tyler, Jeffries & Davies, 1988). Intonation can signal doubt or tentativeness in the teacher’s mind, or indicate whether a student response is expected or not. Loudness and softness can be employed to signal whether the message is for the whole class or a single student.

When teaching online, however, the quality of the voice can be such that much of this subtlety and variability is lost. Due to low quality microphones, poor setups and/or busy internet connections, many teachers struggle to maintain the presence of their voice (Biasutti, Antonini Philippe & Schiavio, 2022). A voice that may be lively and interesting in a F2F class can sound thin and tinny over a Zoom connection or may sound distant and dispassionate. Delays in transmission may mean that questions go unanswered or humour unappreciated. This loss in the quality of voice makes it more difficult for the teacher to create and maintain presence in the online classroom and may lead to students who disengage and find the class ‘boring’. When teachers do encourage their students to speak, the teacher may find that the student has similar or worse voice quality setup issues and the communication is further degraded.

Voice Audio Rating Scale (VARs)

In order to highlight the need for good quality voice audio during online teaching, and to enable students to collaboratively assess each other's voice quality, I have developed the Voice Audio Rating Scale (VARs) which I will outline here and how it can be used in class. The scale is shown below in Figure 2 and links to sample audios can be found by following [this link](#). The scale is not designed to assess pronunciation or articulation but merely the quality of the voice signal in terms of fidelity as it is transmitted from teacher to student over VCS such as Zoom or Teams.

The scale consists of five regular ratings (A to E) and two outliers (S & F). The S outlier is of studio quality and is typically obtained by professional broadcasting and recording companies such as the BBC. It is probably beyond the reach of most teachers since it requires expensive microphones and software to capture the sound of the voice; it is presented here as a gold standard.

RATING		ABSOLUTE	RELATIVE
S	Studio		
A	Excellent		
B	Good		
C	Average		
D	Poor		
E	Failing		
F	Failed		

Zoom jitter?	YES	NO
---------------------	-----	----

Figure 2. Voice Audio Rating Scale (VARs)

The five regular ratings (A to E) represent the quality of voice that teachers typically operate in when online. A and B are green categories and represent the standards that teachers should aim for if they wish to maintain presence in the online classroom and transmit all the subtleties of the segmental and prosodic features of the voice. Ratings C, D and E marked in orange colours represent categories that we should aim to improve. The F category is an outlier and represent a failure to transmit any useful voice that is intelligible.

The VARS scale has two columns, one for *absolute* voice quality and another for *relative* voice quality. Absolute here is the quality achieved given the teacher’s microphone, setup and recording equipment. It would be the quality obtained if a teacher recorded themselves on their PC or laptop and then played it back to themselves. However, when teaching online the voice needs to be transmitted by the VCS over the internet and into the speakers or headphones of the students. The resulting quality is the relative quality, i.e., the quality of voice as perceived by the listeners (the students). The relative quality is a complex outcome of the VCS and setup, the internet connections both for the teacher and the student, and the listening environment of the student. Relative quality is the most pertinent quality for teaching purposes, but teachers can improve this by attending to absolute quality.

The scale also has a checkbox for ‘Zoom jitter’. Zoom jitter (or lag) is when the audio signal breaks up for a moment for the listener which can result in words and phrases being dropped from the speech stream, or delays and compression in packets. Jitter occurs due to a number of complex interacting features including the software and the setup, the upstream and downstream bandwidth and the load on the internet (Chang *et al*, 2021). The speaker is not usually aware of jitter occurring since it is a listener effect. The only way a speaker knows if it is occurring is to ask a colleague to monitor the audio while they are speaking. Since jitter can seriously impair audio quality and intelligibility at all levels on VARS, it is imperative that this check is carried out on a regular basis.

In order to achieve A & B standards in voice audio, the type of microphone, location and setup are all important features that a teacher must consider. Figure 3 shows the quality of each level and the typical setup arrangements on VARS.

Rating	Quality	Typical setup
S	Studio Quality	Professional broadcasting and recording companies such as the BBC.
A	Excellent	Quality USB microphone, close to the mouth.
B	Good	Quality USB microphone but distant.
C	Average	Laptops mics, webcam mics, phone mics, earphone inline mics.
D	Poor	Laptop and webcam mics but poor setup.
E	Failing	Cheap microphones, poor setup and/or background noise.
F	Failed	Unintelligible.

Figure 3. VARS and microphone setups

The microphone is the most important part of the audio chain and the area where a teacher can make the most significant difference in presence. In professional recording studios, expensive microphones are used, maintained and ‘guarded with their life’ since these studios understand the need to capture the voice as cleanly and clearly as possible at source. English language teachers cannot be expected to purchase such systems but should be expected to spend at least £100 (\$120) on a good quality microphone which, if maintained well, should last a career. The microphone should be a USB type which means it has a cable which plugs into the USB port on the PC or laptop. There are many such microphones available but two that I normally recommend are the Blue Yeti by Logitech or the ATR2500x-USB by Audio Technica. These will, in theory, enable teachers to achieve levels A & B on the (absolute) scale. Teachers should not rely on built-in laptop, phone or inline earphone microphones since these are invariably low quality.

The second most important feature is the setup of the microphone in terms of distance to mouth. The quality of the sound from the mouth drops off rapidly with the square of the distance and most experts recommend that the microphone is placed approximately six inches (15 cm) in front of the mouth. This can be checked by extending the thumb and little finger on one hand from the mouth to microphone as shown in Figure 4. Any further away and the quality of the voice captured by the microphone starts to deteriorate, and the advantage of a quality microphone is rapidly lost.



Figure 4. Optimal mouth-to-microphone distance

The difference between the A and B VARS levels in the samples provided on the website is in fact simply the distance from mouth to microphone. The same microphone has been used in each case, but the B sample is approximately twelve inches (30 cm) away from the mouth, a mere six inches more than the A sample. While the B sample is still intelligible, there is a noticeable ‘boxiness’ with the sound resulting from speaking in a sizeable room. The A sample, in contrast, is sharp, upfront and ‘present’ as if the speaker is standing in front of you in the same room. (The term ‘presence’ in professional recording parlance originated from the idea that the sound is close and in the room.)

In practice, we can expect a drop of at least one level on the VARS scale from absolute to relative terms. So a quality microphone that is well setup to produce level A in absolute terms will typically deliver level B in relative terms as shown in Figure 5. In other words, the teacher

produces A but the students hear B. This is the inevitable consequence of the VCS and the internet connection, and to some extent the listening environment the student is in. This fact highlights the need for teachers to insist on producing the excellent A quality standard voice audio since there will inevitably be a loss in quality as the voice travel over the system. In some cases, the drop in level may be even more severe.

A	Excellent		
B	Good		

Figure 5. Typical drop in VARS level from Absolute to Relative

In practice, I find that many teachers and teacher-trainees operate at C levels and below on the VARS scale simply because of the microphone they are using. Many educators rely on the built-in microphones in laptops to pick up the sound, or the microphones located on their webcams. These are rarely high-quality microphones, even if the laptop is expensive. These microphones also restrict their positioning since they are built-in to the device and unmovable. A laptop microphone for example will typically sit 24 inches (60 cm) from the mouth which is excessive and cannot realistically be improved by the teacher leaning in. Mobile phone and in-line earphone microphones also are invariably low-quality as well. While these do enable the positioning of the microphone to be closer to the mouth, the quality of these microphones is such that any gain from positioning is negated by microphone build quality. In the C sample provided, a mobile phone microphone was used. While the sample appears to be loud, the quality of the voice is degraded and the subtle segmental and prosodic features of the voice are likely to be degraded. This sample illustrates the notion that we should not mistake loudness for quality of voice. It is relatively easy to increase the volume of the voice but much more difficult to improve the quality.

Level D ratings on the VARS scale usually occur due to a combination of low-quality microphone and poor setup. In the sample, the teacher is using a laptop microphone. This entails a low-quality microphone which is too far from the vocal source. The result is a thin, boxy sound in which a lot of the subtleties of the voice are lost. Maintaining presence for any length of time in class is virtually impossible at this level of quality.

Level E rating illustrates a low-quality audio voice but with background noise and some jitter. The voice is distant and thin and suffers from noise that masks the message and jitter on part of the message. It is not clear where this noise or jitter originates, and the speaker may not be aware that it is occurring, but the result is a voice quality that is barely intelligible.

The VARS rating scale samples A to E are meant to be indicative of the five levels typically encountered in VCS classrooms. Each level can be perceived through a variety of combinations of microphones and setups. For example, an A level rating might become a B level rating if the microphone is not positioned close to the mouth. An E level rating might become a C or D if the source of noise and jitter is removed. The rating is a subjective rating and works best when a multitude of voices can be heard and compared.

Level F ratings occur when the voice is unintelligible and no useful information can be obtained from it. Level F ratings are invariably due to setup and internet issues rather than cheap microphones and are usually noticed relatively quickly and, hopefully, solved. Without any useful communication between a teacher and students, the class does not get very far. However, finding a solution can take time and a busy teacher may not have this time to attend to a student who is experiencing such problems. (The next section lists a number of steps a user can make to troubleshoot such issues.)

Using VARS

I use VARS regularly in my online classes with teacher-trainees to emphasise the importance of attending to voice quality when preparing to be a teacher and to help the trainees improve their own setup for the class. I start early in the module by introducing the students to the VARS scale and use this as a template for discussing audio matters and choice of microphone. At some point in the module, I will design a task where students rate each other's voice audio in breakout rooms before returning to the main class to summarise as a cohort. I believe it is also important for my own voice audio as a teacher to be assessed by the students. This task serves as a useful prelude to their own preparations if part of the assessment for the module includes student presentations. As part of the feedback for the assessment, I will normally indicate their VARS rating.

It is important to emphasise that using VARS and discussing audio technical features is not simply carried out because we desire to be audiophiles, but because we care about our profession and the classrooms we work in. A lack of attention to audio equipment and setup will result in low voice quality and hence presence in the classroom. Any teacher working at C and below levels of voice audio is seriously impairing their ability to connect with the class and engage in the 'spirit of care, authenticity, immediacy and empathy' that Umpleby (2014, p. 26) demands. Moving online means that more weight is thrown on to the voice in maintaining presence. We cannot leave responsibility for this classroom feature to the IT department to find a solution any more than we would expect them to turn on the computers for us or close a window to dampen outside noise.

Here is a short text that you can ask your students to read the next time they are online in order to check their VARS setup. The text is relatively simple and only takes about 20 seconds to complete meaning that you can complete a whole class in a matter of minutes.

The pet store: George is at the pet store, looking at what kind of pet he might want to get for his birthday. George asked if he could have a horse, but his parents said no because horses are too big. First, he sees dogs and cats. Baby dogs are called puppies. Baby cats are called kittens. George likes them because they are easy to take care of and can play a lot, but they will get bigger. George wants a small pet.

Testing and maintaining voice audio over VCS is not a one-off affair however. I find that regularly checking and testing, and also encouraging students to check each other in breakout rooms, works best. Circumstances can change from class to class as microphones placements and internet strengths vary.

Here is a list of additional checks a user can make to improve connections:

- Move laptop closer to Wi-Fi router.
- Use a wired connection (Ethernet) if possible rather than Wi-Fi.
- Remove any nearby devices which might interfere with Wi-Fi signal (e.g., mobile phone).
- Rotate laptop or PC to obtain better connection.
- Plug laptop into power rather than using battery mode.
- Close other unused applications.
- Switch off HD video in Zoom if not needed. Use standard definition.

(See also Kavanagh, 2022)

Hybrid Teaching

After the full force of the COVID-19 pandemic abated, many teaching institutes went back to teaching face-to-face (F2F) fully in class. Some institutes, including my own, however, have developed hybrid delivery in which a class is taught F2F and online synchronously. This means the tutor is teaching to a cohort of students physically in front of them in class while also running VCS at the same time with a cohort of students online. Hybridity of delivery can occur at the level of the course where some classes are taught fully online and some fully in class, i.e., blended delivery (Nørgård, 2021), but I will not be concerned with this type of delivery here. This article focuses solely on class hybridity (sometimes termed parallel teaching) which can be defined as the synchronous teaching of two sets of students, one online and one in class.

Hybridity has a lot of advantages for an institute. Running classes in hybrid mode means that students can choose to attend physically in class or online depending on their preferred mode of study. Hybridity can also be beneficial for those international students who cannot attend due perhaps to delays in obtaining visas or COVID-19 regulations in their own country. Hybrid classes can also enable a student, who would normally attend physically in class, to study from home if, say, they are sick or test positive for COVID. This flexibility in delivery and mode of study means that institutes that fully grasp the hybrid delivery, I believe, will have a competitive advantage over those institutes that insist on returning to the traditional fully F2F delivery. The attraction for students in the flexibility of delivery that comes with hybridity cannot be overestimated and institutes that embrace this technology are likely to be the gold standards for the future. While a few ‘open’ universities have known this for some time, in today's competitive environment, many traditional universities are also beginning to see the advantages of essentially offering on-campus and distance study modes.

Hybrid delivery of classes is not without problems however. Most teaching institutes and classroom setups have not been designed to combine online and F2F teaching and in many cases teachers have been left up to their own devices to implement ad-hoc solutions in terms of which software to run, where to place cameras and microphones, and how to interact with what is essentially two separate cohorts of students each vying for the teacher's attention at the same time. Setting up a hybrid class at the start of class can be time consuming and technically challenging. Beside preparing for the F2F class, the teacher needs to start up the VCS, admit students and share their screen, all to be done while handling the in-class students.

The main question for teachers delivering lessons in hybrid mode is whether they can maintain presence in both spaces and how to achieve this. The hybrid mode introduces a unique set of circumstances in which the teacher psychologically teaches in two separate and distinct spaces: the physical and digital spaces. As cognitive beings, I do not think we have evolved to handle this situation effectively. For example, with whom do we make eye contact: the students in class or the camera projecting to the students on VCS? Inevitably there is competition for cognitive resources between these two spaces and teachers can feel they are drawn to an imaginary third space which exists just above the eye line of the students in class and just below the eye line of the camera mounted on the ceiling or wherever. This state of affairs is not made any better if the camera is a wide-angled lens designed to take in the whole of the class rather than a facial camera.

When considering the voice, hybrid delivery carries with it many of the pitfalls of fully online mode but has additional complexities which mean that presence will inevitably suffer. First, the microphone provided in the teacher room may not be of the same quality as the one that a teacher has at home or in their office. Most microphones are wide-angled devices that are designed to pick up sounds from all sides of the room which are more suited to business meetings than teaching classrooms. In this case, it is probably more imperative that as teachers we take in our own quality microphones and use USB connections.

Second, the problem of maintaining a close distance from mouth to microphone is exacerbated. The teacher tends to use their space by moving about to create presence with the students in the classroom, but this inevitably takes them away from the microphone, typically located at the front of the class near the PC equipment. This means that the audio quality to the online students is impaired. Staying near the lectern and the microphone is one possible solution but this tends to defeat the purpose of physical classroom teaching which encourages the use of personal space.

One possible solution for this is to wear a lapel microphone which travels with the teacher as they move, but this comes with a series of issues as well. Lapel microphones are rarely of high quality and cannot match the high-spec USB microphones mentioned earlier. They also require connection to the PC via Bluetooth or other wireless connection system. However, this type of connection can suffer from interference and noise, and requires that the microphone be recharged regularly.

Hybrid classes are rarely equally balanced. Typically a teacher might find that early on in a module there are more students in class and fewer online. This situation often reverses itself as the module progresses. Students tend to drift towards online delivery and the number of students physically present in class can dwindle. Undergraduate classes also tend to have more students in class than postgraduate classes, especially when the postgraduate class has a large international contingent. With these asymmetric numbers, a teacher can be drawn to the cohort with the largest numbers. It is very easy to neglect online students if there are only a handful compared with those in class. The online component can often feel like an afterthought. Similarly, classes with only one or two students physically present can seem odd if most of the students are online. What is clear though is that teaching a hybrid class is vastly different to teaching a class solely online. Being physically in class and having to handle two set of students in different spaces is quite different to sitting at a desk in front of a PC or laptop and speaking to a cohort of students wholly online. The interesting cognitive challenge of hybrid delivery is something that may take time for teachers to accommodate to.

Class discussion can also suffer in hybrid delivery. If a teacher asks a question, this may be picked up by a student in class, but will not be broadcast well to the online cohort, especially if the student is distant from the teacher's microphone. This may lead to the online cohort sensing that they do not have a voice in the class. The delay in transmission further means that the in-class students have advantages when it comes to interacting with the teacher. One solution is to have a second microphone which can be handed round the classroom so that the student's voice is picked up and relayed fully to the online cohort. However, this not only means the teacher needs to charge and setup another relatively expensive microphone, but it often entails that the class discussion is turned into a moderated affair in which the teacher has to relay every answer and idea from one cohort to the other. Students themselves often do not have the requisite skills to communicate effectively between spaces and thus the teachers find themselves in the situation represented by Figure 6 whereby no pathway between the two sets of students can be reliably established.

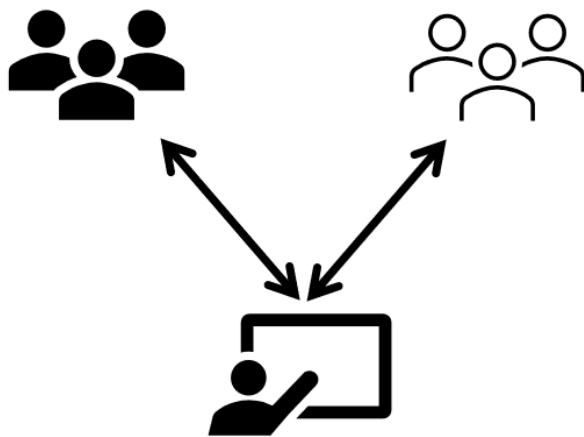


Figure 6. Hybrid teaching scenario

Conclusion

With the move to fully online and hybrid teaching, the need for teachers to maintain presence with their students has never been more pressing. Traditional communication through physical presence, body language and space is being superseded with a heavy emphasis on the voice as the sole carrier of information. As teachers, we need to understand how our voices are being transmitted through and across the internet into our students' rooms if we are to maintain, and even enhance, the care, authenticity and empathy (Umpleby, 2014) we want to give our students.

The technical challenge of getting the best voice audio over VCS is not an easy nut to crack so to speak. We will need to spend time and effort in this endeavour and become familiar with a field that we did not think was ours. But as teachers, we do need to do this ourselves since no one else will. We cannot rely on our institute managers or the IT department to setup our voices any more than we expect them to close the windows or turn on the computers in the classroom for us. The call here is not simply a call to become audiophiles for the sake of it. It will be teachers who suffer through a loss of presence, unmotivated students and frustrating classes if we do not learn the technical details of voice audio transmission so that it becomes second nature to us.

As we go forward, teaching circumstances will no doubt change. New VCS will come on the market and different ways to interact with and teach students will appear. Already, augmented and virtual reality systems are enabling educators to interact in 3-D digital spaces with their students to create immersive learning experiences that far outweigh the experience of 2-D VCS (Meccawy, 2022). But the teacher's presence in these environments will always be the driving force of these experiences, and the human voice will always be at the heart of this connection, enabling and facilitating the care, authenticity and empathy our students deserve.

About the Author

Michael Cribb holds a PhD from the University of Reading in Applied Linguistics. He taught language and linguistics in the Far East for many years before returning to the UK in 2004. He taught at Oxford Brookes before moving to Coventry in 2007 where he is currently Assistant Professor in English Language. His teaching and research includes phonology & grammar, psycholinguistics, AI & mind, and discourse analysis. He has published several books including "It's Language, Stupid" – a journey into the DNA of the mind.

To Cite this Article

Cribb, M. (2023). The power of the voice in facilitating and maintaining online presence in the era of Zoom and Teams. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, <https://doi.org/10.55593/ej.27106int>

References

- Biasutti, M., Antonini Philippe, R., & Schiavio, A. (2022). Assessing teachers' perspectives on giving music lessons remotely during the COVID-19 lockdown period. *Musicae Scientiae*, 26(3), 585–603. <https://doi.org/10.1177/1029864921996033>
- Chang, H., Varvello, M., Hao, F., & Mukherjee, S. (2021) Can you see me now?: a measurement study of Zoom, Webex, and Meet. IMC '21: Proceedings of the 21st ACM Internet Measurement Conference, November 2021, 216–228. <https://doi.org/10.1145/3487552.3487847>
- Halonen, J. S. (2002). Classroom presence. In W. Buskist & S. F. Davis (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer*. Psychology Press.
- Hardy, S. (2021). Teacher presence: What is it and how can you achieve it? *Twinkl Digest*, <https://www.twinkl.com/news/teacher-presence>
- Hincks, R. (2005). Measures and perceptions of liveliness in student oral presentation speech: A proposal for an automatic feedback mechanism. *System*, 33(4), 575-591. <https://doi.org/10.1016/j.system.2005.04.002>
- Kavanagh, J. (2022). Hacks and tips to improve Zoom call quality. *Jotform*. <https://www.jotform.com/blog/zoom-video-quality/>
- Lehman, R. M., & Conceição, S. C. O. (2010). *Creating a sense of presence in online teaching: How to "be there" for distance learners*. Jossey-Bass.
- Mahendru, H. C. (2014) Quick review of human speech production mechanism. *International Journal of Engineering Research and Development* 9(10), 48-54.

- Meccawy, M. (2022). Creating an immersive XR learning experience: A roadmap for educators. *Electronics*, 11(21): 3547. <https://doi.org/10.3390/electronics11213547>
- Moorhouse, B. L., Li, Y., & Walsh, S. (2021). E-Classroom interactional competencies: Mediating and assisting language learning during synchronous online lessons. *RELC Journal*, online Feb. 20, 2021. <https://doi.org/10.1177/0033688220985274>
- Newton, R. (2019). *Authentic gravitas; Who stands out and why*. Seven Dials.
- Nørgård, R. (2021). Theorising hybrid lifelong learning. *British Journal of Educational Technology*, 52(4), 1709-1723. <https://doi.org/10.1111/bjet.13121>
- Port, R. (2007). The graphical basis of phones and phonemes. In O. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 349-365). J. Benjamins Pub.
- Su, A., & Wilkins, M. (2013). *Own the room: Discover your signature voice to master your leadership presence*. Harvard Business Review Press.
- Tyler, A. Jefferies, A. A., & Davies, C. E. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teacher's spoken discourse. *World Englishes*, 7(2), 101-110. <https://doi.org/10.1111/j.1467-971X.1988.tb00223.x>
- Umpleby, G. (2014). *Towards an understanding of presence in teaching: Having and being*. PhD Thesis, University of Exeter. <https://ore.exeter.ac.uk/repository/bitstream/handle/10871/14938/UmplebyG.pdf>
- Wells, J., Butt, F., Orsini-Jones, M., Borthwick, K., Preshous, A., Chen Y., Ngoc, T. B. T., Nguyen, H., Nguyen, T. T. T., Nguyen, A. D. (2022). *A ViVEXELT (Viet Nam Virtual EXchange for English Language Teaching) guide for E-CIIC mediators in synchronous online zoom breakout rooms*. ViVEXELT, vivexelt.com. https://vivexelt.files.wordpress.com/2022/04/a-vivexelt-guide-for-e-ciic-mediators_final.pdf