Enhancing Online Language Education Through Virtual Reality

Jason TACKER

The purpose of this research is to review problems of video conferencing language lessons and propose an additional enhancement through using virtual reality. Video conferencing is a simulation of in-person human interaction, attempting to imitate the rhythms of face-to-face conversations, gestures, symbolism, identities, and body language. The lack or non-existence of many of these in-person education aspects creates difficulties for language classes. Therefore, this research will first review and explain the problems of video-conferenced classes. Next, it will point out how virtual reality can address these problems, or the lack of multimodality (using multiple literacies to understand one medium) by demonstrating in virtual reality step by step how participants create their own identities, use virtual bodies or "avatars" for body language, gesturing, and creating spaces and props to better symbolize authentic classrooms.

Keywords: motivation, classrooms, Virtual Reality, EFL, identity, video conferencing

In the 1970's The Japanese roboticist Masahiro Mori at the Tokyo Institute of Technology coined the term "uncanny valley". He created this term for the different types of negative feelings brought on by the viewing or interacting with robots. People are willing to accept that robots are human-like up to a point, however beyond that is too close to "human" and creates negative feelings, or feelings of unease in people which in turns means they lose their affinity for robots (MacDorman, 2017).

Video conferencing is its own version of the uncanny valley. The programs we now widely use for communication such as Zoom are simulations of true interactions, but not quite right. This video conferenced "uncanny valley" depletes the energy of its users and often creates negative feelings and unease.

In Blum's (2020) article *Why we're Exhausted by Zoom* she states that there are reasons for the feelings of exhaustion. Zoom works best in an authoritarian lecture format with a leader in control of the situation, formal rules and structure are adhered to and everything runs smoothly. However, in more interactive classrooms that want autonomous learning in a non-authoritarian style such as ESL classes, all the tools of human interaction are needed.

For non-authoritarian, student-centric and autonomous classes to work efficiently, multimodality must be an integral part of the system of the class. Multimodality is the use of multiple literacies in one medium (Kress, 2010). Such literacies of a class would be

similar to Georgia Tech's definition of multimodality for its Communications Department written, oral, visual, electronic, and nonverbal. (https://wcprogram.lmc.gatech.edu/guiding-principles/woven) These are all "modes" or cooperative tools with which everyone in the class use to communicate with each other.

In a face-to-face classroom, teachers and students have a greater perception and agency of the tools of multimodality. Head nodding, distractions, cues that encourage speakers to begin or end, eye gazing, signifying practices are beyond the oral, written, visual and electronic modes of multimodality that can be completely lacking or misunderstood as they are projected to a screen and separated by squares in video conferencing. When people look down in video conferenced class, are they looking at the class text or the game they are playing on their smart phone? When we see the student looking down, we want to know what it means, it is human nature to ask the question: "Why did they do that?" (Blum, 2020).

Many of the modes of face-to-face communication, something we find natural and easy to do, are far more difficult to decipher in video conferencing and take a great deal more effort. Regardless humans will naturally attempt to interpret them anyway, to exhaustion.

Due to the rapid increase in video conferencing, the director of the Stanford Virtual Human Interaction Lab (VHIL), Jeremy Bailenson (2021), has been studying specific psychological causes of what is now being called "Zoom fatigue" due to its over whelming popularity in use in recent history. They systematically deconstructed Zoom fatigue and assessed Zoom on its individual technical aspects. What they found were four consequences of prolonged video chats contribute to the feelings of exhaustion.

To Close for Comfort

Bailenson (2021) states that the amount of eye contact and the size of the faces on screen is unnatural. In normal face to face meetings, gatherings, classes or public speaking situations, people will be looking to various points of interest, writing notes down, checking mail on their phones as a few examples. In contrast, on zoom everyone is looking at everyone all the time. The dynamic of listener and speaker is not traded off as in normal communication and everyone is treated like a speaker all at the same time. It is similar to one of the biggest phobias that exist, public speaking. The eye contact is dramatically increased and that creates a stressful experience.

To a lesser extent but still prevalent can be the faces sizes of people in a video conference. Depending on your screen size, or if connected to an external monitor, human faces can appear larger than comfort allows. In one-on-one conversations through conferencing or Zoom is switched to a single presenter mode, if the face is close up to the screen, you are seeing the persons face at a size that simulates a personal space that is normally associated with intimacy. When someone's face is that close to ours, the brain interprets it as an intense situation that is either means mating or fighting (Bailenson, 2021).

The Parakeet Effect

Stated plainly, looking at your image in the mirrored video feed for long periods of time is fatiguing. Most video conferencing programs have a square feed that shows you what you look like, it is an incredibly unnatural way to communicate.

In the real world, if somebody was following you around with a mirror constantly - so that while you were talking to people, making decisions, giving feedback, getting feedback - you were seeing yourself in that mirror, that would just be crazy. Nobody would ever consider that. (Ramachandran, 2021, para. 11)

Many people are now seeing themselves for many hours every day. It is taxing and stressful with many negative emotional consequences, as psychological research has shown that seeing your reflection, you become far more critical of yourself (Bailenson, 2021).

No Place to go, nowhere to run

A large portion of human interaction, communication is committed to while on the go or during some larger form of movement. In person and phone conversations give people the autonomy to move around or walk if they choose to. With video conferencing a camera has a specific field of view and participants are stuck within that box for the duration of the dialogue, thus movement is limited in ways that are not natural to people. Bailenson states There are several studies showing that locomotion and other movements cause better performance in meetings. For example, people who are walking, even when it is indoors, come up with more creative ideas than people who are sitting (Oppezzo & Schwartz, 2014). Much of that work shows a causal relationship-for example, children who are required to gesture with their hands while learning math showed more learning retention compared to a control group (Cook et al., 2008).

There is an illusion, Bailenson says that we have but impossible to have in video conferencing. He gives an example of talking on the phone with what he imagines the demeanor of the listener to be. He imagines that the listener is giving him his unfettered, undivided attention even though He is not reciprocating. He is multitasking on many various sub activities as is natural yet continues to believe that his listener is not doing the same but is his captive audience. He believes this will be the downfall of video conferencing, with the illusion of a captive audience for our dialogues gone, we will lose the ability to multitask making video conferencing more of a burden. (Bailenson, 2021 Para. 29)

Cognitive Baggage

Bailenson states that in reality our non-verbal communication flows naturally without much thought put into our gestures or nonverbal cues. Research has shown (Kendon, 1970) that nonverbal behavior is effortless and still complex when in synchronous communication. In video conferencing the behaviors are still effortless, but our interpretation of those behaviors has become much harder.

In Zoom for example it is much harder to send cues. Behaviors are monitored and created intentionally to send generated messages, examples are centering oneself in the cameras field of vision, exaggerated movement like head nodding to show agreement, or directly looking into the camera instead of faces on screen to show eye contact. Participants also raise their voices by 15 percent (Cores et al, 2019) when on video conference calls.

The receiving end of communication can also be hard to interpret. In a face-to-face conversation, people draw great meaning from head and eye movements, which help to signal turn-taking, agreement, and a host of affective cues (Kleinke, 1986). What happens when these cues are present and perceived by other communicators but are not tied to the intention of the person making the gesture?

Zoom users face this disconnect often. For example, in a face-to-face meeting, a quick, sidelong glance where one person darts their eyes to another has a social meaning, and a third person watching this exchange likely encodes this meaning. In Zoom, a user might see a pattern in which on their grid it seems like one person glanced at another. However, that is not what actually happened, since people often don't have the same grids. Even if the grids were kept constant, it is far more likely the glancing person just got a calendar reminder on their screen or a chat message. (Bailenson, 2021, para. 21)

During a video conference, participants are constantly getting a stream of non-verbal cues that would have specific meanings in reality but take on completely different meanings in video conferencing.

All of these, sending and receiving of cues and other non-verbal behavior, take more effort than natural flowing conversation and can become cognitively taxing on communicators (Bailenson, 2021).

Identity is Motivation

Blum (2020) and Bailenson (2021) have intersecting points in video conferenced classes. For Blum it is the lack of multimodality, and for Bailenson it is the exhaustion from interpretation of sending and receiving modified cues from witnessing the new forms of behaviors adherent to video conferencing. Both have come to the conclusion that what was easily interpreted in the past in face-to-face classes has now become muddled and confused with interpretations, cues and behaviors that can't be accurately made without causing problems such as "zoom fatigue" or exhaustion. There is a far more serious problem associated with this as extensive loss of clear and easy interpretation can lead to a decline in identity and motivation in students.

After a series of video conferenced classes, it was by chance of a harmless rhetorical question that the feeling of not being students was brought to light in the paper *Video Conferencing Effects on Student Culture and Motivation in L2 Classrooms*. The instructor had asked students if they enjoyed zoom classes, and not to his surprise he received a res ounding "NO" to his query. Upon asking why they did not like it, the answer surprised him, noting that they did not feel like university students going to a university (Tacker, 2021).

Students were trying to interpret the behaviors, new signifying practices of video conferencing and cues just to come up short in comparison with the ideals of what they had

already known or thought that their university lives would be like.

During Zoom classes students did not have to adhere to the culturally or socially acceptable norms of dressing for school, or going to a place called "university" along with all of its behaviors, such as eating in a cafeteria, going to classrooms, attending clubs or events, using specific sub-culture language, Etc. In many cases they did not need to even use the video camera or microphone during video conferenced classes taking away even more ability to understand the cues, something that would not happen in a face-to-face class where existence in class is an automatic; everyone sees each other and interact easily. This exhaustion furthered the thought processes of students, they did not really feel like they were committing to the actions of real students and thus in some ways they did not have confidence in their situation as being university students (Tacker, 2021).

This new formed conception of lacking the identity of a student as they understood it caused them to lose motivation in classes. This is best understood through the Identity Based Motivation Theory. IBM states that a person has a preconceived conception of what their identity is that they are thinking about in any given situation. When behaviors are undertaken in a specific context and actively thought about, the individual judges if those behaviors are identity congruent (student behaviors for example). If the behaviors are determined to be incongruent with the pre-conceived identity of that context (What students believe a "student" is), then the behaviors, cues, and signifying practices lack meaning and there is no motivation to continue them (Oyserman, 2014).

The inherent problems of video conferencing that are proposed by the above researchers can be summed up to a deficiency of something. Lack of meaning, of signifying, symbolizing, cue understanding and multimodality. The continuous lacking these important aspects is not congruent with what teachers and students believe their identities to be.

All people expect that we have specific places, behaviors, cues and practices that create the identity that we need to feel motivated to deal with the situation at any given time. For students and teachers, our cues and behaviors must be of the appropriate identity congruent context, otherwise we lose motivation over time to continue them. Video conferencing, such as Zoom does exactly that; it can demotivate learners when the congruence between actions taken and the identity presumed do not match in the context.

Virtual Reality: A Potential Tool for Distance Learning

Virtual reality is the use of computer systems to simulate an environment for human interaction. Different from the keyboard and mouse, VR makes use of a head mounted display (HMD) with which users see the virtual world that they are interacting with. VR can simulate many senses at the same time: vision, touch, hearing and in some experimental cases, smell. VR is only limited by its availability of content and the processing power of the computers used.

Currently the most popular, commercially available HMDs are from the Facebook owned Oculus brand. The Quest series of HMDs are both a self-contained VR experience provider but also can be linked to a PC with graphics processing power for greater graphic fidelity and content outside of the Oculus software library.

Among the numerous types of software are virtual chat rooms, software for meeting

people from all over the world to talk and participate in multiple activities beyond just chatting. One of the most popular of these chat rooms is *Rec room* (https://recroom.com/), a multi-room simulation of a recreation room and dorm similar to a hostel or student hotel. Users of Rec room have their own dorm room and are invited to make their own creative use of the program free of charge.

The most useful and relevant way that Rec room has been implemented is for education. Checking through YouTube, an interested educator could find videos such as "*Area of Composite Figures - Area of Triangles and Rectangles Lesson*" (https://www.youtube.com/watch?v=0JZ8BXfXGzI) or "Language exchange in VR" (https://www.youtube.com/watch?v=tx8rKy1_0Ac). Both videos give examples of how versatile VR simulations can be, how it could be applied to language courses and that VR is possible for distance learning for ESL students at a university.
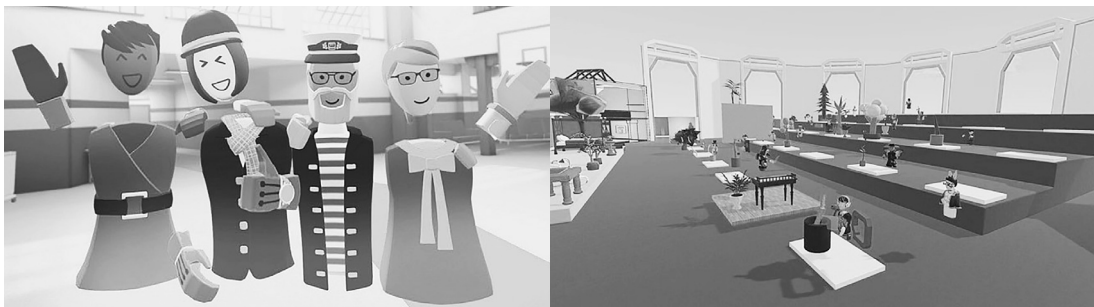
Virtual Reality for Motivation and Identity

Almost all of what is lacking or exhausting from video conferencing software such as Zoom can be regained through the use of virtual reality for education. There are two key components that are necessary to accomplish this: a virtual body and a virtual space.

In Rec room and all other VR applications the user has an *avatar*, or a computer-generated image of the user that they manipulate in the virtual world. Depending on the software and HMD this avatar or VR body can have all the anatomy of a normal person. Typically, avatars lack legs only; in reality we are not looking at our legs as we walk or run but rather straight ahead. We more often see our arms and hands in a first-person fashion.

The other component is the virtual space. The virtual space is where avatars interact and can be anything that users want the spaces to be. If a person wants to make a classroom space, the tools are easy to understand and often free of charge within the larger virtual reality chat software. A space is limited only to what the user desires and what limitations the VR application may have. It is ideal to note here that while most VR software is "free to play" , in this sense you can download and use the software for free, even build a space or customize an avatar to your liking; however, they have premium services that users can pay for. Such services are special types of clothing for avatars, hairstyles, skin colors accessories, and for spaces how much space a person can use to build in and how many objects they can place in the virtual space.

With both a virtual body and a space to use for classroom instruction, a teacher and stu-

Example- Rec room avatars and a classroom with object creation lesson (Rec room, 2020).

dents have everything to combat exhaustion, create motivation and identity.

A virtual avatar is an extension of the self in the online world. A user can walk, run, move their arms, or head. A person can speak into the VR microphone and their virtual mouth moves, face expresses, and other VR users can see and hear them. They can use their hands to manipulate objects in sophisticated ways. Clothing can be interchangeable, facial hair even gender depending on how a user feels about themselves at the time.

The virtual space or room is whatever is desired giving the same community context that is needed for all users involved to understand their roles within the space. As a classroom would have a functioning white board with markers, desks and multimedia experiences for teachers to use for facilitating lessons, so also does the VR space. Demonstrations with props can be accomplished easily for a lesson. One example could be teaching specific language for immigration: simulating an airport checkpoint, with costumes, passports and even metal detectors.

Virtual reality is a greater form of representation of reality than a video conferenced square on a screen. Eye contact is not constrained to the flat screen but can roam over an entire environment giving learners a more natural way to look around or make meaningful eye contact with others. Each learner does not see a mirror box of themselves but views from the first person in a realistic way. Behaviors, context, cues and presence are virtual but understandable in a way that is human and as natural as the everyday world. A VR environment gives a more natural sense of space as learners can move around being as close or as far from others as they wish. These all contribute to creating the virtual identity of learners, making them actually feel like students thus motivating them, because they can signify through their avatars and the space they inhabit. The ability to express themselves in virtual reality as they choose to by taking advantage of VR's ability to mimic reality also lessens the cognitive load because the tools for expression of non-verbal behavior are present and almost as natural to interpret as the real world.

Conclusion: Is VR feasible? Is it applicable?

As one might expect from a new technology, it has its benefits but also its problems. As it stands the biggest concern for most interested parties such as universities, price being the biggest of those.

In a previous section, it was mentioned that some virtual reality HMDs can run software stand alone or for better quality experiences can be connected to powerful PCs. The only HMDs that can accomplish this are from the "Oculus Quest" series from Facebook. They come in two forms: CV1 which is the older Quest 1 headset, and the CV2 the newer version. Both of these headsets run in the few hundred-dollar range, depending on how much system memory you want for storage.

The other type of HMDs are the PC only connected headsets. These headsets must be connected to a powerful PC to run only PC virtual reality software from systems such as Steam, Oculus, or WMR (Windows Mixed Reality). These more premium of headsets range from seven hundred dollars for the standard WMR HP Reverb G2 to the VIVE Pro Eye at nearly two thousand dollars.

It is easy to understand that for both universities and students the price for admission

could very well be a sticking point in the adoption of virtual reality for classroom use. It would cost a university quiet a sum to supply each student with a VR set or even be a small financial burden on a student if they had to buy one to attend virtual classes. But, if the price tag is not a problem, recent research suggests that application of virtual reality in classrooms can have motivating effects and lower exhaustion.

In a paper on the Bibliometric analysis of Virtual reality over the last two decades, research indicates that the United States, United Kingdom and Taipei are the top 3 most productive countries/regions which are involved in virtual reality research in education (Liu, et al., 2017). The majority of research done on virtual reality has been done primarily in English speaking countries and with the exception of Taipei, no other Asian countries are working as seriously on virtual reality which leads to a belief that this could be a cutting edge form of Educational research in Japan.

In the field of English as a Second Language or English as a Foreign Language, there is an ever-growing amount of research on the effects of VR on ESL or EFL classrooms. One such paper used the ImmersMe VR platform in the university's department of languages and culture to test English speaking ability from level 1 to 3, or from beginner to advanced. They found that VR was an ideal platform for enhancing the different skills of EFL because of its variety of contexts (multimodality) to learn communication skills from and its immersiveness for students (Bendeck, et al., 2020).

Another example of VR research in ESL and EFL was conducted with high school students. The researchers created a spherical video-based VR system for English-speaking training with a peer assessment approach. They found that the VR approach improved English-speaking performance, motivation and critical thinking (Shu-Yun, et al., 2020). As well they found that VR reduced the student's English learning anxieties. Students were able to understand behaviors and cues easily and felt less exhaustion.

This final article researches the effects of virtual reality on PBL, or Problem Based Learning for specialized English, such as English for Engineering purposes. Their research examines how the integration of VR technology into PBL contexts affects students' motivation for, problem-solving during, and vocabulary acquisition in learning English as a foreign language (EFL). What they found was that VR significantly enhanced the learning of specialized vocabulary, enhanced the motivation and sustained the EFL student's interest in English and enhanced the participant's problem-solving performance (Chen, et al., 2021).

Ultimately virtual reality is a growing field that has proven to have beneficial uses in EFL education for academic purposes. The ability to adapt a virtual world and people to any specialized language situation can both reduce or eliminate the fatiguing use of video conferencing completely, create identities that are congruent for students and motivate them in many more ways than even a traditional classroom can. Currently even with proof of the benefits of virtual reality from the above sources in Asia, many students, parents, and universities may find the difficulties such as temporary discomfort or expense of HMDs to difficult to hurdle.

References

Bailenson, J. N. (2021). Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. Technology, Mind, and Behavior, 2(1). https://doi.org/10.1037/tmb0000030

Bendeck Soto, J., Toro Ocampo, D., Beltrán Colon, L. & Valencia Oropesa, A. (2020). *Perceptions of ImmerseMe Virtual Reality Platform to Improve English Communicative Skills in Higher Education.* International Association of Online Engineering.

Blum, S.D. (2020). Why We're Exhausted by Zoom Inside Higher Ed. https://www.insidehighered.com/advice/2020/04/22/professor-explores-why-zoom-classes-deplete-her-energy-opinion

Chen, C-H, Hung, H-T, Yeh, H-C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. Journal of Computer Assisted Learning, 37: 851- 860

Ho, C.-C., & MacDorman, K. F. (2017). Measuring the Uncanny Valley Effect. International Journal of Social Robotics, 9(1), 129-139. https://doi.org/10.1007/s12369-016-0380-9

Kendon, A. (1970). Movement coordination in social interaction: Some examples described.Acta Psychologica, Amsterdam, 32(2), 101-125. https://doi.org/10.1016/0001-6918(70)90094-6

Kleinke, C. L. (1986). Gaze and eye contact: A research review. Psychological Bulletin, 100(1), 78-100. https://doi.org/10.1037/0033-2909.100.1.78

Kress, G. R. (2010). Multimodality: A social semiotic approach to contemporary communication. London: Routledge.

Liu D., Bhagat K.K., Gao Y., Chang TW., Huang R. (2017) *The Potentials and Trends of Virtual Reality in Education.* In: Liu D., Dede C., Huang R., Richards J. (eds) Virtual, Augmented, and Mixed Realities in Education. Smart Computing and Intelligence. Springer, Singapore

MeCey, B. (2017) *VR and 'Rec Room' Are Changing the Way We Socialize* https://virtualrealitypop.com/vr-and-rec-room-are-changing-the-way-we-socialize-55f6d1a7bd71 (Image)

Oyserman, D. (2014), "Identity-Based Motivation: Core Processes and Intervention Examples", Motivational Interventions (Advances in Motivation and Achievement, Vol. 18), Emerald Group Publishing Limited, pp. 213-242. https://doi.org/10.1108/S0749-742320140000018006

Ramachandran. V. (2021) Stanford Researchers Identify Four Causes for 'Zoom Fatigue' and Their Simple Fixes. Stanford News https://news.stanford.edu/2021/02/23/four-causes-zoom-fatigue-solutions/

Recroom (n.d.) *Player Created Creative Classes!* https://recroom.com/creative (Image)

Shu-Yun Chien, Gwo-Jen Hwang, Morris Siu-Yung Jong, (2020) *Effects of peer assessmentwithin the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions*, Computers & Education, Volume 146, 103751

Tacker, J, (2021) *Video conferencing effects on student culture and motivation in L2 English Classrooms*, Akita Prefectural University RECCS Bulletin, Volume 22, Pg. 59-66