Lijing Shi* and Ursula Stickler

Eyetracking a meeting of minds: teachers' and students' joint attention during synchronous online language tutorials

<https://doi.org/10.1515/jccall-2021-2006>

Received May 20, 2021; accepted July 23, 2021; published online September 6, 2021

Abstract: The COVID-19 pandemic forced many language teachers to move their teaching online without sufficient preparation. This unexpected change of practice engendered doubt and anxiety in teachers. They worried about their ability to attract and hold their learners' attention, an element that is essential for successful online teaching. Our dual-point eyetracking study looks into how students and teachers establish joint attention during online language tutorials. It also examines various means teachers employ to guide students' attention and scaffold their meaning making process. The data was collected from two online language tutorials where the eye movement of one teacher and one student was tracked simultaneously, as well as recordings of their stimulated reflection while watching their own eyetracking visualisation replay. By combining mixed-method data and dual perspectives, we were able to unveil the complex interactions in online language tutorials and offer practical suggestions to language practitioners who hope to improve their online teaching skills.

Keywords: eyetracking; joint attention focus; online language teaching; pedagogy; SCMC

1 Introduction

As experienced online language teachers and researchers, we witnessed the wave of online teaching forced by COVID-19 lockdowns with mixed feelings. It has been an unusual year, in which the previously non-mainstream practice of online language teaching suddenly became the default. It is exciting to see how most language teachers took to online teaching, expanding their teaching arena and developing new teaching skills which will be essential for their future practice.

***Corresponding author: Lijing Shi**, London School of Economics and Political Science, London, UK, E-mail: L.Shi3@lse.ac.uk

Ursula Stickler, The Open University, Milton Keynes, UK, E-mail: ursula.stickler@open.ac.uk

Meanwhile, we also heard of recurring issues and difficulties reported by these novice online language teachers. As a majority of them were forced into online teaching without choice, prior experience, sufficient training, or profound understanding, their workload and anxiety soared (Talidong & Toquero, 2020). This is not surprising, considering that fewer than half of all teachers felt well prepared for online teaching before the pandemic according to an OECD TALIS survey (OECD, 2019). One of the biggest sources of anxieties came from teachers not knowing whether their teaching was working in an online environment, or in other words, whether they successfully directed students' attention during online lessons (Gao & Zhang, 2020; Stickler & Shi, 2015). All the non-linguistic cues teachers used to rely on in a face-to-face class – such as eye-contact, gestures, nods – became unavailable or at the least less reliable and timely; not to mention that many teachers looked at the black box of students' video window and asked themselves: “Am I talking into a void?”

Rather than seeing online language teaching as a less than ideal replacement for face-to-face teaching, teachers need to be aware that online communication is ontologically and epistemologically different. Another layer of mediation is added to the teaching process. “The mediating effect of online communication technologies means that the ways in which learners and teachers make meaning and create inter-thinking spaces are different compared to face-to-face environments” (Stickler & Hampel, 2019: 19). Such awareness is particularly important for language teaching, because, there, communication is not only the means of teaching but also the goal. Effective online language teaching is essentially dialogic. Teachers do not merely aim to transmit knowledge, but also need to enable learners to internalise it through interaction during a dialogue, and furthermore to interact effectively in order to internalise this knowledge in the target language. Such interaction requires both sides to pay attention to salient points (e.g. content, form, feedback) at the right moment.

The concept of joint attention originates in studies of child language development (Tomasello & Farrar, 1986) where episodes of joint attention are linked to heightened linguistic activity in the child, and subsequently to larger vocabulary gain. While we cannot directly transfer insights from child learning processes to adult language learners, joint attention has been recognised as essential for social learning (Kim & Mundy, 2012). Methods to investigate joint attention in adults include eyetracking (Caruana et al., 2018; Pfeiffer, Vogeley, & Schilbach, 2013).

As mentioned above, the attention focus is essential for language learning processes. Online environments allow for synchronous or asynchronous communication and teaching. Joint attention can be achieved in online synchronous communication (SCMC) through not just “simultaneous occupancy of the communicative space” but also “both individual and joint (mutually known)

attention to unfolding meaning” (O’Rourke & Stickler, 2017, p. 3). Online language teachers therefore need to effectively guide learners’ attention throughout the lesson as well as paying attention to learners’ responses and needs. More colloquially, we could describe this successful online language teaching as a “meeting of minds”. In order to successfully establish joint attention through technology-mediated online language tutorials, we need to better understand what is happening on both sides of the communication: we need to analyse the attention focus of both, teachers and students. This study, therefore, will search for evidence from eyetracking that the attention focus of teachers and students aligns during online language tutorials, it will also aim at reassuring teachers that their attempts to direct their students’ attention can be successful, and – finally – it will identify some successful means that teachers employ to achieve joint attention.

This paper first presents online language teaching from a sociocultural perspective, and argues that research methods need to be adjusted to deal with new questions in the field. The reasons for employing a mixed method dual-point eyetracking approach are discussed, followed by an exploratory, proof-of-concept study involving two online language tutorials. By examining qualitative data from eyetracking visualisation and stimulated recall reflections, this paper demonstrates the evidence of a “meeting of minds” and key themes relevant to online language teaching.

2 Literature review

2.1 A sociocultural perspective of online language teaching

In sociocultural theory, the most fundamental concept is that “the human mind is mediated” (Lantolf, 2000: 1). Human beings do not react directly to physical environments, but respond to cultural and social artifacts (e.g. signs, tools) to make meaning and establish connections. From a sociocultural perspective, learning is dialogic and interactive. It takes place through dialogues between learners and their more able peers (including their teachers). Therefore, learning is socially situated and knowledge is jointly constructed. Knowledge creation is an ongoing process of “teaching-and-learning”, and its effectiveness depends on the quality of such dialogues. Learning seen as a cultural process means that knowledge is not possessed individually but co-constructed amongst the members of communities; this co-construction can be supported through scaffolding.

Online language learning, as well, has to be “dialogic” as traditional lecture-style, transmissive teaching cannot enable a learner to master the target language. A learner needs to work with input, raise questions, practise and receive feedback until knowledge and skills are fully internalised. During this dialogic learning

process, the interaction between a language teacher and a learner is a dialogue in the first (L1) and the target (L2) language, influenced by individual sociocultural features. Furthermore, online language learning is different from face-to-face language teaching, as technology adds another layer of mediation due to cultural backgrounds of its creators. In addition to the sociocultural elements of teachers and learners brought to play in the teaching dialogues, the process and the product of such dialogues are mediated by the affordances of the technology.

Hampel holds that “[m]ediation is a crucial concept in sociocultural theory, where interaction is seen as crucial to learning” (Hampel, 2014: 90). To successfully establish a dialogue in synchronous online communication, we need joint attention, i.e. an agreement on the focus, content and boundaries of the communicative space (O’Rourke & Stickler, 2017). Online interaction with regards to language teaching is based on the joint attention of the teacher and students on salient elements. Online environments take away some cues which teachers rely on to confirm that “attention is met” (e.g. eye contact, body language). Therefore, teachers feel more confused, less confident or increasingly anxious: Do online learners actually look at the places they are directed to pay attention to? Are they paying attention to something as expected from their teacher? If so, how can teachers be assured that their expectations are met?

According to O’Rourke (2008) four aspects of online language learning have been neglected, one of which is attentional focus, i.e., what users are actually attending to at a given moment (2008: 233). O’Rourke and Stickler (2017) point out that “joint attention is the key property” of synchronous communication (SC), because more than simultaneous presence in the same space needs to be achieved: a reassurance that attention on making sense together, focusing on the same issues helps to create SC (O’Rourke & Stickler, 2017).

According to Shvarts and Abrahamson (2019), “the construct of joint attention captures the phenomenon of two (or more) people focused on the same object and each aware of the other’s focus”. This concept has inspired much research on early childhood language acquisition, the development of social skills and other cognitive functions. However, research on the process of establishing joint attention in SC has been rare. This is largely due to the difficulties of employing the right research tools and methods. In the section below, we are going to propose an innovative way of advancing research in this area.

2.2 Combined dual-point eyetracking with stimulated reflection

As discussed earlier, online language tutorials are mediated by technology, and the way how learners and teachers make meaning and interact is distinct from

face-to-face classrooms. This requires us to innovate our research method, moving from quantitative methods to a mixed method approach, moving from only telling one side of the story (teacher or students' behaviour, intention, attention, etc.) to embrace both sides' perspectives, ideally both at the same time. Our approach is to combine dual-point simultaneous eyetracking with follow-up stimulated reflection from a teacher and a student who participate in the same online synchronous tutorial.

In the last few years, we have carried out multiple eyetracking experiments within a sociocultural paradigm (Shi, Stickler, & Lloyd, 2017; Stickler & Shi, 2015, 2017), to ascertain online participants' attention focus patterns and the rationale identified by them. Diverging from traditional eyetracking research, which is heavily based on statistical data, we combined eyetracking data with qualitative data collected from follow-up stimulated reflection interviews when the participants watched the recordings of their own eye movements during online language tutorials. This mixed method approach has proven to be fruitful, because it (1) Helped to reveal the pattern of online participants' attention focus, (2) Triangulated the reasons for devoting attention to specific areas, and (3) Enhanced participants' reflectivity and awareness of their teaching and learning behaviours. For example, our eyetracking study on online Chinese learners (Stickler & Shi, 2015) discovered that all beginners looked at Pinyin during a reading task regardless of their linguistic levels. The follow-up stimulated recall interview revealed that the reasons for using Pinyin ranged from meaning comprehension, to consolidation or confirmation of characters and Pinyin. During an interactive task, 70% of learner's attention was on content, 20% on social areas of the interface, and 10% on technical information. Participants also verbalised their needs to pay attention to peers' social presence.

Using the same research method, we were able to uncover the intricate dynamics between online teachers' attention focus, teaching skills and online teaching experience (Shi et al., 2017). It was found that teachers' attention to the content, social and technical areas is influenced by their familiarity with the online teaching tool as well as their prior online teaching experience. A highly experienced online teacher who was familiar with the tools could focus more on the teaching content and social areas without paying much attention to the technical area. If a teacher was unfamiliar with the specific online teaching tool or had little experience of online teaching, the attention focus on the technical area was significantly higher. Compared to the student study, this study found teachers' attention on social areas slightly higher than the proportion of students' attention on the same areas. The reason was that teachers were clearly aware of the limitations of online communication and were trying to compensate for the loss of non-verbal cues by paying attention to students' social presence indicators

(e.g., emoticons in the participant window). The stimulated reflection data revealed that more experienced tutors spent more time on attending to social and content areas and reflected more frequently about online teaching skills at a higher level, based on a comparison with Hampel and Stickler's pyramid of online teaching skills (Hampel & Stickler, 2005).

However, the above two studies investigated the perspectives of learners and teachers separately, making comparisons only in retrospect. As we are interested in joint attention between online teachers and students synchronously, the next step is to track both sides' eye movements simultaneously, i.e., use dual-point eyetracking. To date, there have only been a small number of dual-point eyetracking studies related to online education.

Belenkuy et al. (2014) conducted a dual-point eyetracking study to explore primary students' joint attention patterns when they were collaboratively solving mathematical problems online. This study followed a typical eyetracking research convention attempting to link eyetracking data with pre-test and post-test results to establish statistical correlations. By comparing "gaze recurrence plots" of high-performing and low-performing dyads on solving procedural learning problems, they discovered that the high-performing dyads appeared to have a high degree of joint attention. They also found that joint attention was related to learning gains in conceptual knowledge in the procedural condition.

To verify Vygotsky's learning theory, Shvarts and Abrahamson (2019) explored teaching/learning collaboration when the student solved challenging maths problems. They monitored the perceptual activity of four teacher–student dyads by enlisting two eyetrackers and having the dyads sit in front of the computer next to each other in the lab. They were able to offer an empirical operationalization of the teaching/learning process which is "dynamic transformation of intersubjective coupling between the student and tutor perception–action systems" (Shvarts & Abrahamson, 2019).

These two dual-point eyetracking studies strengthened our confidence in expanding our mixed method approach to involve both the online teacher and student in the eyetracking study at the same time. The above studies have proven that dual eyetracking is a valid method for investigating concepts like joint attention, collaboration and interaction which all play crucial roles in synchronous online teaching and learning. However, both studies used a dual-point system where both participants were co-present in the same room and also followed a conventional eyetracking analysis approach without bringing in participants' reflection. Creating a novel dual-point eyetracking method where teacher and

learner are physically separated, simulating true distance learning with Synchronous Computer Mediated Communication (SCMC), as well as employing our proven sociocultural method of merging eyetracking data with stimulated reflection, will allow us to develop more ecologically valid accounts of how joint attention is established.

Taking a sociocultural perspective, where shared understanding forms the basis for learning, the current study will combine teacher and student perspectives, tracking two participants in online language tutorials simultaneously, to uncover how the student's attention shifts in different parts of online tutorials, or how students may struggle to focus on the relevant area due to technical, linguistic or other difficulties. It will explore the answers to the following research questions:

- 1) Do learners and teachers successfully establish joint attention during synchronous online language tutorials?
- 2) What means do online teachers employ to guide the attention of their learners to specific aspects of the online task?

Sub-question for 2): Which of these means are successful and which are unsuccessful?

3 Research procedure

The experiment consisted of two steps – (1) simultaneously tracking teachers' and students' eye movements during the same online tutorial, (2) following this up immediately with a stimulated reflection session where both parties were invited to watch their eyetracking visualisation videos in parallel and to reflect on their teaching and learning behaviours, with special emphasis on their attention focus and underlying reasons.

This study was carried out under lab conditions at the Open University, where the teacher's eye movements were recorded on a table mounted Tobii 60 eye tracker (62.5 Hz sampling rate). A mobile eye tracker (Tobii Pro X3-120, 120 Hz sampling rate) attached to a laptop was used to track the student's eye movement.

Two short online language tutorials for adult learners were pre-designed. Two language teachers and two language students were invited to the university lab in order to be eye-tracked on site. The other students participated online at a distance. Due to ethical restrictions, all participants were also language teachers learning an additional language.

Before the experiment, the researchers had set both screens used for the online tutorials and eyetracking identical in dimension and interface to ensure Areas of

Interest (AOI) would be the same for both sides. Both the teacher and the learner went through nine-point calibration to ensure the accuracy of eyetracking data. When all the preparation was ready, the online tutorial and the eyetracking started simultaneously.

The first tutorial was for German beginners: one German teacher (T1) taught four students for 15 min. One of the students (S1) was invited to participate in another lab room and his eye movements were tracked by a mobile Tobii eye-tracker. Once the tutorial ended, both eyetracking recordings stopped at the same time. Then both T1 and S1 were invited into the observation room where they could view their eye-movement gaze-plot visualisations replay on side-by-side wall-mounted screens. The teacher and the student watched the recordings in parallel, commenting as they went along. Two researchers sat behind them and recorded the dialogue. Once the replay finished, the researchers joined the conversation by commenting and asking questions to encourage more reflection from both sides. There were no pre-designed or semi-structured questions.

The second tutorial was an English (ESL) tutorial for intermediate students. There were four adult learners and one teacher (T2). During this 15-min tutorial, the main objective was to learn the word order of adjectives in English. After the teacher’s instruction, students needed to re-order a sentence with different adjectives by dragging and dropping the textboxes on the screen. The eye-movements of the second teacher (T2) were tracked as well as those of the second student (S2). The same research procedure as for the German tutorial was followed. Table 1 summarises the background information of these two tutorials.

Table 1: Background information of two online tutorials.

Language	German	ESL
Level	Beginner	Intermediate
Content	Likes and dislikes	The order of adjectives in sentences
Teacher	Teacher 1: Female, native German speaker, experienced online teacher	Teacher 2: Male, native English speaker, experienced online teacher
Students	Student 1 and three other students (J, L, T).	Student 2 and three other students (J, L, E).
	Student 1: English in his 60s.	Student 2: European in her 40s.
	Student J: South Asian in her 20s.	Student J: South Asian in her 20s.
	Student L: European in her 40s.	Student L: European in her 40s.
	Student T: English in his 50s.	Student E: European in her 40s.
Duration	15 min	15 min

4 Data and findings

To give a fuller picture of how students and teachers interacted and to what extent their attention focus met or missed, this section will first present the eyetracking visualisation data of the main teaching phase and then scrutinise representative stimulated reflection excerpts.

4.1 Eye-tracking visualisation data

Two tutorials took place via OU Live, a synchronous online teaching tool. As shown in Figure 1, the screen can be divided into two parts – the microphone button, participant window, textchat and editing tools were on the left, and the whiteboard was on the right. In the case of the German tutorial, the whiteboard was divided into top and bottom sections. T1 presented key grammar points on the top. The bottom space was later used to display students' own sentences during the interactive speaking task. The teacher used “crosses” and “smiles” to present different degrees of likes and dislikes in German.

Figure 1: Interface, gazeplot and ‘hand’ pointer of the German online tutorial.

Figure 2 shows the interface from a student’s point of view later in the tutorial session when students started to produce their own sentences utilising the grammar structure they just learned.

Figure 2: Interface from a student's point of view.

When T1's and S1's eyetracking recordings started to replay side by side, it was apparent to all participants that both parties' attention met. Figure 3 displays four pictures of combined attention focus during two phases of the task in the German tutorial. It shows the student's gazeplots on the left, and the teacher's on the right.

The top two images, Figure 3a and b, illustrate the eye movements of the student and the teacher at the beginning of the task when the teacher was explaining a grammar point in German. The teacher's gazeplot moved from technical aspects (e.g., microphone button) to the content on the screen. The student's gazeplot shows his gaze was following the teacher's gaze almost

Figure 3: (a) (top-left): Student's gazeplot during instruction. (b) (top-right): Teacher's gazeplot during instruction. (c) (bottom-left): Student's gazeplot during interactive speaking. (d) (bottom-right): Teachers' gazeplot during interactive speaking.

instantly. When the teacher was explaining the grammar point, his attention was on the content, largely overlapping.

Figure 3c and d respectively illustrate the student's and teacher's gazeplot patterns during the phase of the task when students were using newly learned grammar structures to make their own sentences one by one. There is a visible increase of the teacher's attention on the participant window and textchat box. This is a clear contrast to her attention focus pattern during the instruction phase when she was mainly focusing on the content. This shift in attention might be indicating that she was planning the next question to ask or thinking about inviting the next student to speak.

Similarly, the student's attention focus became more complex as well. His main attention was on the content. Attention to the top half of the screen indicates that the teacher's sample structures were used as scaffolding to support his understanding and language production. Strikingly, when his peers spoke their own sentences using suggested structures, much of his attention was on the structures listed at the lower part of the whiteboard. This provides evidence of "learning from peers" in a live online session and exemplifies the idea of "knowledge is co-constructed". He also focused on the names in the participants window when his peers were speaking. Such behaviour is linked to social presence issues. He also looked at the microphone icon without which his voice would not be heard.

Next, Section 4.2 will summarise two stimulated reflection interviews, and Section 4.3 will discuss the reflection data according to the six most relevant themes.

4.2 Stimulated reflection data summary

The first reflection interview started from Student 1 (S1) explaining his eye focus, which was on the microphone button at the very beginning of the tutorial. Teacher 1 (T1) immediately noticed the pattern of her eye movements (which were constantly on textchat and participant window). Then, both of them witnessed the "meeting of minds" from the parallel eyetracking recording re-play. The recordings show that S1's gaze followed the teacher's pointer, which was in the shape of a hand. They continued with comments on each other's eye movements, attention focus, and learning/teaching behaviour. More evidence of the teacher's skills was gradually unveiled, and so were her thoughts behind those behaviours. Both parties also discussed an incident caused by a technical issue which was also related to her particular online teaching strategy.

The second interview, reflecting on the ESL session, started with both parties (T2 and S2) confirming that their eye movements "met", looking at the same area, more or less at the same time. They then commented on their online behaviours

from two sides. T2 found out that he had missed the information in the chat due to changes of the position of the elements in the software interface. T2 also had a chance to hear that his less precise instruction caused students confusion. S2 discovered that she did not only follow the teacher's instruction, but also paid attention to peer social presence. This recall session ended with T2's reflection on online teaching in general – how students reacted differently in different tasks, the impact of different tools, and what he would have done differently.

The following section will examine six key themes in online language tutorials against the data from these two groups' reflection stimulated by watching their eye movement recordings in parallel.

4.3 Recurring themes during online tutorials

4.3.1 Meeting of minds – joint attention

Playing the eyetracking recordings side by side generated ample evidence that students' attention focus synced with their teachers'. For example, T1 commented on how the student's eye movement went hand in hand with her, by saying: "That's amazing how parallel our eyes go" (T1: [00:03:22]).

The data show that students' attention was guided by the teacher's verbal explanation (Extract 1) as well as by the teacher's use of the pointer tool (Extract 2).

Extract 1 T1: [00:07:01] Then I said 'verb', and you looked at the 'verb'!
 T1: [00:07:21] Again you look that's so reassuring. You're just follow, that's parallel.

Extract 2 S2 [00:01:18] I'm reading the sentences and as soon as your pointer moves, I go up again to the pointer.

By having both the student and teacher involved in the reflection at the same time, it became possible to have one side's observation and reflection confirmed by the other side simultaneously. This confirmation, reassurance and triangulation are only possible by employing a complex mixed-methods research methodology – such as our innovative set-up. Extract 3 is an example of how both parties confirm that the student's attention was guided by the 'hand' (the icon the teacher chose here for the pointer tool). Furthermore, the student was fully aware of what to pay attention to – not the symbol of the hand-shape cursor, but the learning content the 'hand' pointed to. Such discovery should give practitioners more confidence that students do pay attention to the information their teachers try to provide.

- Extract 3 S1: [00:02:07] Yes. I follow it.
 T1: [00:02:09] Yeah, exactly. Oh, you follow it. Oh yeah..... (*She laughs*) So, that is interesting. So, students do! You do! You follow my cursor!
 S1: [00:02:18] Because I was very aware. I thought, no, I don't follow the hand, I follow what the hand is pointing to
 T1: That's amazing!!
 S1: [00:02:30] So, so we should see me looking to what you're pointing at, not to the hand itself.
 T1: [00:02:37] Yeah, exactly to what yeah, absolutely, but in a way, you followed my 'hand'. That is interesting! I always feel it's important to use the 'hand'. Yes. Yeah. The pointer.

The above data supplements and confirms the eyetracking visualisation data in Section 4.2.

4.3.2 Online teaching strategies

The eyetracking data reveal that teachers paid attention to the content, technical and social areas during online tutorials. Generally speaking, when a teacher explains a grammar or language point, his/her attention is predominantly on the content area, paying occasional attention to the technical and social areas. When students start to interact either with the teacher or with their peers, the teacher's attention focus spreads out to the social area and technical information as well.

During the reflection session, T1 noticed her own way of teaching took the shape of a triangle – moving from content to participant window to textchat box and back to content. A considerable amount of her attention was on the social area especially when students practised their own sentences. The list of names in the participant window was a virtual presentation of her students. When she was worried about a weaker student, J (name abbreviated for anonymity), she kept checking the textchat box and participant window anticipating that the student might 'raise her hand', or enter comments or questions via the textchat. Extract 4 is an account of how a language teacher manages her online teaching – not just focusing on the content she needs to teach, but also in preparation of different responses from her students, being fully aware of students' understanding levels and difficulties, as well as paying attention to potential reactions from them. This is a typical example of what a highly skilled online teacher attends to during a live tutorial.

Extract 4 T1: [00:03:] That's amazing how parallel our eyes go. It's just my eye, obviously, as you say, it is my job, goes more often from time to time to the left, checking, checking the textchat and also the participant window. I was a bit worried about J*.

Another interesting discovery (Extract 5) was that the teacher provided further scaffolding opportunities for weaker students. In her group, one student (J) was a weaker student in terms of linguistic skills, so T1 always left J as the last one to speak in a sequence so that the student had more time to prepare and also an opportunity to learn from the students before her.

Extract 5 S1: [00:14:38] You went down and then you went back up.
 Researcher 1: [00:14:40] It was always T*, Ch*, L*.....
 S1: [00:14:43] You either top to bottom or bottom to top. You never went randomly.
 [...]
 T1: So, the way I went, you at the top, I started with you. I always left out J*. I want to know, is that T* or even started with T*. And J* always last, so she had enough practice.
 S1: [00:15:35] I still thought I was going to be first.
 T1: [00:15:37] Yeah, I thought you were once or twice first. But I think I like because I knew, as a tutor, you know your students; and I knew T* is confident and that's why I asked him first.

4.3.3 Misalignment and miscommunication

As discussed earlier, an online language tutorial is a complex and interactive process mediated by technology. It is common for mismatches of attention and expectation to take place. For example, in the second tutorial, the teacher (T2) did not pick up students' questions typed in the textchat, because the textchat box was presented in an unfamiliar screen position. This small incident reminds us that the interface of the online teaching environment matters. Teachers need to be familiar with the tools and features of the online teaching platform, as well as students' communication channels. In a sense, online teaching is more demanding than its face-to-face equivalent as teachers need to attend to multimodal information within a very short span of time.

Extract 6 records another misalignment caused by interface differences. The "arrow" sign was displayed differently for T2 "the icon with the arrow" and S2 "the arrow", which resulted in S2's confusion when she heard 'the icon with the arrow'.

This echoes with Hampel's discussion (2019) about the impact a different interface and computer setting could have on synchronous online teaching.

Extract 6 S2: [00:03:09] – which is funny because when I started, I already had the arrow. And then you confused me by saying “you need to click on this, on the icon with the arrow”. Well, I’ve already dragged something. I don’t know if you could see that I dragged the whole frame, the yellow frame around

However, not all miscommunication or misalignment leads to failure. Here (Extract 7) is an example of how an experienced teacher dealt with misalignment which reflects her teaching beliefs. When T1 could not hear S1's verbal responses due to technical issue (S1's microphone was switched off by someone else), she asked twice and then decided to move on. In the recall session, she explained that the reason was to save S1's face and lower the student's anxiety, to avoid “putting the student on the spot”.

Extract 7 T1: [00:12:26] The reason I do this, is because, I went on I didn't want to have you too long on the spot. I asked once or twice and then maybe the student has decided not to speak, or it's too difficult. I just go on then. That's why. So you don't feel on the spot. Everybody is waiting.

Anxiety is a common theme in language education, particularly in speaking. Online students' anxiety level can go even higher because of technical issues. It is important for teachers to consider their students' emotional wellbeing and try their best to construct a safe and welcoming online learning environment.

4.3.4 Paying attention to social presence

We view learning as essentially a sociocultural activity and the same is true for online language learning. Our eyetracking data reveal evidence of students paying attention to other students; consequently, we can assume that this attention can also lead to learning from their peers. Our reflection data can provide proof that this is the case. Extracts 8 to 13 from the reflection interviews show some examples of this. Despite the fact that they were participants at a distance, their attention focus, from time to time, was on the list of names in the participant window and on information in the textchat. This is very similar to students looking at their peers, or glancing at their facial expressions in a face-to-face classroom.

- Extract 8 S1: [00:02:52] And there..... I was just being nosy and seeing what other people were typing because one of the participants couldn't answer.
- Extract 9 T1: [00:04:29] What is interesting, you still focus on the words while T* was talking and others speaking, you don't look in the participants, but I do obviously.
S1: [00:04:40] Well, I did occasionally
- Extract 10 S1: [00:05:09] And again, I am being nosy, I am looking at the textchat to see
- Extract 11 T1: [00:05:29] And you look all over the place there. Whilst I was explaining this.
S1: [00:05:38] Curiosity.
- Extract 12 S2: [00:02:12] -and now I'm looking at T* and sort of cheating and seeing if he can
T2: And then he went and changed it.
- Extract 13 S2: [00:05:10] – And here again, you are checking the correct order, and I'm sort of just,
T2: [00:05:16] – You know, looking over other people's shoulders.

Extract 14 records that S2 noticed T2 using an emoticon and commented on it. From the reflection we know that T2 received confirmation from students about using emoticons for online teaching. In addition to top-down advice from the university, witnessing, by watching the eyetracking recording and hearing from S2, that students actually pay attention to such teacher-organised social presence had a very positive and reassuring impact on his teaching.

- Extract 14 S2: [00:13:11] I do actually notice that you gave an applause here.
T2: [00:13:18] – that's a good point when the student..... We tell colleagues when we train them to use OU Life and make a big thing about it use the emoticons, persuade the students to use emoticons and that to develop that cohesion with your group. I wonder if the students notice them.
S2: [00:13:35] They do, they do
T2: [00:13:42] You've got evidence that they look over and see them?
Researcher: always, always, in the participants window there is about 20% attention in there.
S2: [00:13:48] They do look for them.

4.3.5 Anticipation and preparation

Apart from paying attention to teacher's instruction and peers' online social presence, students also displayed an interesting trait – anticipating what is going to happen. Students mentally prepare a question they think might come up. In the reflection session T1 noticed that S1's attention moved to the grammar structure he would be asked to use ("You're getting ready with your eyes?"). S1 explained that he figured out which question he would be asked and was preparing for it, instead of paying attention to the sentence T1 was explaining (Extract 15).

- Extract 15 T1: [00:09:03] Do you know – you can see yourself thinking. you looked at the two smiley faces, then you went up there, you look what is two smiling faces and you checked it (zigzag) back. Yes. Yeah, it's your turn. You keep checking
 S1: [00:09:28] Once you asked me and you couldn't hear me. And I thought I took the microphone off.
 T1: [00:09:39] That's you now.
 S1: [00:09:40] I did this one.
 T1: [00:09:43] You're getting ready with your eyes.
 S1: [00:09:45] Yes, kind of.
 T1: [00:09:51] One. Yeah. Yeah. While you as a student look up to make sure you can construct a sentence. I, I've just seen, I look just on there, on the bottom part and the participants.
 S1: [00:10:07] There's also in the face-to-face classroom, if you go down in order, the students know which question they're going to be asked and they switch off. Yes, I switched off. I was looking at my question, not the one you were currently doing.

This extract has uncovered another important factor in synchronous online teaching, apart from the teacher's intention and students' perceptions: like in a face-to-face classroom, anticipation also plays a role. Shvarts and Abrahamson (2019) point out that strong anticipation is one of "the key mechanisms" in the building up of joint attention. The reason is that "each participant relies on the immediate perception of the shared *living* space as this perception reveals the affordances for action and tunes into the intentions and dispositions of the other participants' actions directly apparent in their gestures and movements" (Shvarts & Abrahamson, 2019).

Anticipation is two-way as well. A teacher anticipates different actions students might take due to their different levels and different needs. For a successful

online interaction, the data suggests that it is not enough for both sides to be fully concentrated, but they also need to be fully “anticipating”. In other words, teaching online is more than being responsive: it has to be proactive and based on a good understanding of students’ needs and thoughts.

4.3.6 Eyetracking as a tool for empowering learning and teaching

Traditional eyetracking studies focus on quantitative data and treat participants (objectively) as informants. The uniqueness of our study is to use eyetracking to support participants’ reflectivity. Data here show that both the teachers and students deepened their understanding of online learning and teaching by watching both parties’ attention focus and commenting on those behaviours.

For example, Extract 1 also recorded the teacher’s relief and excitement when she witnessed how her use of a “hand” pointer successfully guided the student’s attention online. She commented with great delight and vivaciousness on this confirmation of her professional success. Believing in one’s teaching technique is one thing, but having it confirmed with visible evidence engenders a higher level of confidence. This would not have been possible without the evidence from this eyetracking experiment.

There is also verbal evidence that teachers deepened their understanding of students’ behaviour online, highlighting crucial areas for causing potential mismatches of expectation and interpretation. In an afterthought to the reflection interview, T2 reflected on the use of different tasks and how students responded differently because of task differences (Extract 16).

- Extract 16 T2: [00:17:29] We didn’t plan it, but it was quite useful that yours was spoken. They had to speak. (Yes, it was. Yeah, exactly. Yeah. Yeah.) and didn’t have to use the tools. Whereas in mine, they didn’t need to speak at all but they were engaged in a completely different way.
- T1: [00:18:03] It’s just in one activity.
- T2: [00:18:06] in a 90 min, of course, it would be both..... [...] one is speaking, and one is drag and drop.

Observing what has taken place during the reflection sessions, one of the researchers commented, “That was so good. I was nearly moved to tears. It was so fantastic to have you two chatting about it. It was so amazing. You are going to get so much out of that” (Researcher 2).

5 Discussion

This mixed method dual-point eyetracking study followed teachers' and learners' attention focus in the same online learning event. Through reflections stimulated by the video replay of participants' gaze focus, it established where attention focus runs in parallel between the learners and their teachers and where it diverged. This methodological innovation, combining the teacher's perspective with the learners' view, has confirmed where misunderstandings or instances of incomprehension originate, what strategies were used, and where skilful intervention resolved these misalignments and helped to re-establish shared comprehension.

For the first research question, the answer is a simple "yes" as evidenced in eyetracking visualisation and confirmed by the stimulated reflection data. By watching the gaze plot videos side by side, parallel eye movement can be observed, particularly for the instruction part. Where the students engage in interactive tasks, the teacher's eye focus could also encompass social and technical areas.

During the instruction phase, joint attention is established by the teacher directing the students to focus on forms. Present–Practice–Produce is a common language teaching approach even for online lessons. When a teacher is presenting language structures, his/her focus is on the forms, and so is the students who follow teachers' guidance. Our eyetracking visualisation data proves this. In the practice part, the teacher's attention is also on the feedback and on the technical aspects (namely, the social and technical areas) as the teacher's role changes from an instructor to a facilitator. Meanwhile, a considerable amount of students' attention is on peer social presence which once again reminds us that learning is a sociocultural activity. The uniqueness of online teaching allows us to make this joint attention visible.

Conducting this eyetracking experiment with subsequent reflection revealed what joint attention means in online tutorials and how important it is for online teachers to be confident in their ability to establish it. Traditional quantitative eyetracking – only based on the quantitative data – is too limited for researching online interactive teaching. As the teacher's exclamation ("That's so reassuring") confirms, to be able to capture overlapping gaze movements and joint attention opens up a new horizon for researching synchronous online teaching. Based on almost 10 years of experience in conducting mixed method, sociocultural eyetracking, we managed to adapt the research method to fit these new questions. It is necessary to collect evidence for a successful establishment of joint attention in SCMC (O'Rourke & Stickler, 2017), but it is also beneficial for teachers to realise that their methods of online teaching are working; simply because online tutorials are fundamentally different, not just a deficient replacement of face-to-face teaching.

Data shows that students' attention was guided by the teacher's pointer and by his or her verbal explanation. Joint attention, however, goes beyond following the teacher's pointer or looking at where the teacher is looking at that moment, but requires "individual and joint (mutually known) attention to unfolding meaning" (O'Rourke & Stickler, 2017, p. 3). In this case, focusing on what the teacher is referring to. As confirmed by the reflection data, the teacher displayed a wide range of skills to direct a student's attention. For the second question, the answer is that teachers use multimodal affordances offered by the tools available. Past research (Stickler & Shi, 2013, 2017) has shown that teachers who successfully use tools such as the pointer, highlighter, emoticons, raised-hands and textchat as well as verbal communication are more likely to guide learners' attention. Rather than trying to re-create a deficient face-to-face environment online, these successful online teachers expand on the affordances of online spaces, similar to what Stickler and Hampel suggest: "Many teachers as well as researchers continue to judge online learning on the basis of the affordances of face-to-face environments, rather than exploring the additional affordances that the new digital media offer and using them to best effect" (Stickler & Hampel, 2019: 20).

The main pedagogical implications of our findings are that teachers need to develop multimodal competence to fully unlock the potential of a variety of online tools to establish joint attention. For example, during online teaching, teachers have a wider range of tools to give confirmation: in addition to verbal praises like the one in a traditional classroom, the teacher can also use emoticons or textchat to acknowledge students' effort and performance (Stickler & Shi, 2013).

Although the majority of the means the teachers employed in our study were successful in guiding students' attention, it is worth pointing out some unsuccessful examples as we can learn from these how misalignment can negatively affect students and how we can improve the training for online teaching. One unsuccessful incident in the study revealed how, from the student's point of view, questions in the textchat to the teacher were overlooked because of the different positioning of the textchat box on the teacher's screen. Similar to Xiaomei in our 2017 study (Stickler & Shi, 2017), this teacher was very familiar with online teaching, however, the different interface posed difficulty to his teaching. Teachers need to be aware of the interface differences between teacher's view and student's view. In some extreme cases the interfaces for the teacher and students are different because of different permission set by the software or the institution. The interface difference can also be a result of personal choice, for example, if the computer or device is set to a different language, font, or colour. If a teacher explains some features from her/his point of view, she/he needs to be aware that it might look different from students' perspective. Not only for this reason, it is necessary to offer student training or a short session to make sure students know

the common terms teachers use when referring to online tools (Heiser, Stickler, & Furnborough, 2013).

A similar incident where the student shifted some of the shapes on the whiteboard in an attempt to contribute non-verbally revealed how the teacher not acknowledging this contribution affected the student's confidence. An online language teacher who employs different tools and modalities throughout the tutorial needs to be incredibly alert to not overlook any student contribution. For this reason, we recommend that tutors reflect on their own ability and level of skill to create a positive environment for learners while, at the same time, not creating a stressful situation for themselves.

Data in this study thus echoes the Online Teaching Skills Pyramid proposed by Hampel and Stickler (2005). The more experience of online teaching a language teacher has, the more s/he can notice during the teaching (e.g., paying attention to multiple areas), and even plan in advance (e.g., anticipating students' reaction), and thus online teaching becomes an art of directing students' attention at the right time to the right place.

Other pedagogic implications for online teaching derived from this study include: to avoid information overload, for example, by clearing up the whiteboard from time to time; to make teaching instructions more explicit; and to have a pre-announcement of tasks or a tightly structured teaching plan that may lead to less confusion and better teaching results.

Last but not least, one of our future plans is to recommend the use of eye-tracking visualisations to train teachers who are not directly involved in the data collection or have no access to eyetracking equipment. It seems to be a promising avenue for trainee teachers or novice online teachers to observe what took place during a successful online language tutorial by watching such eye movement recordings combined with reflections.

6 Conclusions and limitations

To a great extent, teachers' anxiety and uncertainty about teaching online come from seeing it as –

a limited and limiting endeavour, with computer-mediated communication lacking the depth of face-to-face interaction, offering reduced modalities and not allowing for certain non-verbal and paralinguistic features; it is seen as not immediate, creating cognitive as well as affective challenges (Stickler & Hampel, 2019: 20).

Online tutorials pose a challenge for both teachers and students as communication is mediated by technology. Whether full understanding has been

achieved can only be surmised through deliberately communicated reactions, e.g., ticks or verbal confirmation. This complexity of communication is yet intensified by teaching a foreign language online: the students have to cope with linguistic difficulties on top of the potential technical and media problems. To ascertain whether teacher and students are communicating successfully, following the same trail of thought and paying attention to the same issues, eyetracking has been a valuable method. Recording the gaze focus of learners and teachers in online tutorials separately (Shi et al., 2017; Stickler & Shi, 2015; Stickler & Shi, 2017) has already brought insights into online learning and teaching strategies and potential pitfalls for online instruction.

In this innovative experiment, we expanded our research to include two perspectives in a single learning event: the perspective of a teacher and her/his attention focus with that of a student in one synchronous online language tutorial. In this proof-of-concept experiment two separate eye-trackers were used simultaneously to record and align data derived from both sides. The subsequent stimulated reflection interviews took the format of side-by-side viewing of gazeplot visualisation replay while commenting on the stimulus, effectively creating small-scale focus groups. This method allows the student and teacher to compare and contrast their experience and to support each other in reflecting on the tutorials.

This study combines the perspectives of teacher and student, and tracks two sides in an online tutorial at the same time. The results show the dynamics of establishing teacher-student joint attention; evidencing, for example, how students' attention shifts following a teacher's instructions if the communication is successful, and how the learner struggles to focus on relevant screen areas when the instructions are lost in either technical or linguistic difficulties. This study also provides information behind joint attention, miscommunication and misalignment during online synchronous lessons.

In contrast to other dual-point eyetracking studies, our experiment aimed at recreating the natural environment of online tutorial by physically separating the teacher and student. This allows us to draw conclusions about online communication in the absence of face-to-face cues or a physical shared space. Researchers will gain insights into guided noticing for language learning and the skills employed by teachers and learners to confirm mutual comprehension. The resulting protocol for dual-point or even multi-point eyetracking can be of use for future research in all contexts where shared understanding for the purpose of making meaning online is investigated. In terms of methodology, this project challenges the traditional approach of eyetracking as located in psycholinguistic research (Abdel Latif, 2019) and adds the dimension of learner and teacher reflection by using stimulated reflection interviews with participants, which is cutting-edge in the field.

Despite its significance, our study has obvious limitations in the sense that it is not a typical eyetracking study supported by statistical data and quantitative analysis, and the number of research participants is small. This exploratory study follows a sociocultural paradigm and explores the use of eyetracking as a tool to empower language learners and teachers by encouraging reflection and dialogue. We do not claim our findings as universal, but explore what would otherwise be difficult to see in online teaching without this research method.

The huge increase in online teaching caused by COVID will surely change the practice and landscape of future language teaching, which makes it all the more important that online teaching is conducted competently and successfully. This research has advanced our knowledge of what exactly happens during online tutorials, and can inform future training, not only for language teachers, but for teachers of all subjects where online teaching involves a joint construction of knowledge.

Acknowledgements: The authors would like to thank Dr. Mair E. Lloyd for her contribution in this project as well as all the colleagues who participated in the two online tutorials, in particular Chris Nibblett and Elke St John.

Research funding: This project was funded by CREET, The Centre of Research in Education and Educational Technology at the Open University, UK.

References

- Abdel Latif, M. M. (2019). Eye-tracking in recent L2 learner process research: A review of areas, issues, and methodological approaches. *System*, 83, 25–35.
- Belenkuy, D., Ringenber, M., Olsen, J., Alevan, V., & Rummel, N. (2014). Using dual eye-tracking to evaluate students' collaboration with an intelligent tutoring system for elementary-level fractions. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 36). Retrieved from <https://escholarship.org/uc/item/8m3543hd>.2014.
- Caruana, F., Gerbella, M., Avanzini, P., Gozzo, F., Pelliccia, V., Mai, R., & Rizzolatti, G. (2018). Motor and emotional behaviours elicited by electrical stimulation of the human cingulate cortex. *Brain*, 141(10), 3035–3051.
- Gao, L. X., & Zhang, L. J. (2020). Teacher learning in difficult times: Examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.549653>.
- Hampel, R. (2014). Making meaning online: Computer-mediated communication for language learning. In A. Peti-Stantić, & M. Stanojević (Eds.), *Language as Information. Proceedings from the CALS Conference 2012* (pp. 89–106). Frankfurt am Main: Peter Lang.
- Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, 18(4), 311–326.
- Heiser, S., Stickler, U., & Furnborough, C. (2013). Student training in the use of an online synchronous conferencing tool. *CALICO Journal*, 30(2), 226–251.

- Kim, K., & Mundy, P. (2012). Joint attention, social-cognition, and recognition memory in adults. *Frontiers in Human Neuroscience*, 6, 172.
- Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1–26). Oxford: Oxford University Press.
- OECD. (2019). *TALIS 2018 results: Teachers and School Leaders as lifelong learners (volume I)*. Paris: OECD Publishing.
- O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer mediated communication and language learning. *Computer Assisted Language Learning*, 21(3), 227–251.
- O'Rourke, B., & Stickler, U. (2017). Synchronous communication technologies for language learning: Promise and challenges in research and pedagogy. *Language Learning in Higher Education*, 7(1), 1–20.
- Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews*, 37(10), 2516–2528.
- Shi, L., Stickler, U., & Lloyd, M. E. (2017). The interplay between attention, experience and skills in online language teaching. *Language Learning in Higher Education*, 7(1), 205–208.
- Shvarts, A., & Abrahamson, D. (2019). Dual-eye-tracking Vygotsky: A microgenetic account of a teaching/learning collaboration in an embodied-interaction technological tutorial for mathematics. *Learning, Culture and Social Interaction*, 22, 100316.
- Stickler, U., & Hampel, R. (2019). Qualitative research in online language learning: What can it do? *International Journal of Computer-Assisted Language Learning and Teaching*, 9(3), 14–28.
- Stickler, U., & Shi, L. (2013). Supporting Chinese speaking skills online. *System*, 41(1), 50–69.
- Stickler, U., & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO Journal*, 32(1), 52–81.
- Stickler, U., & Shi, L. (2017). Eyetracking methodology in SCMC: A tool for empowering learning and teaching. *ReCALL*, 29(2), 160–177.
- Talidong, K. J. B., & Toquero, C. M. D. (2020). Philippine teachers' practices to deal with anxiety amid COVID-19. *Journal of Loss & Trauma*, 25(6–7), 573–579.
- Tomasello, M., & Farrar, M. (1986). Joint attention and early language. *Child Development*, 57(6), 1454–1463.

Bionotes

Lijing Shi

London School of Economics and Political Science, London, UK

L.Shi3@lse.ac.uk

Lijing Shi (MEd, PhD) is a multi-award winning teacher at LSE, UK. Her research publications focus on eyetracking, online language teaching, dynamic assessment and intercultural communication. She is a reviewer for several leading research journals. In September 2020, she was elected as the Chair of British Chinese Language Teaching Society.

Ursula Stickler

The Open University, Milton Keynes, UK

ursula.stickler@open.ac.uk

Ursula Stickler is Senior Lecturer in the School of Languages and Applied Linguistics at the Open University, UK, a distance teaching institution. Her research deals with technology enhanced language learning and teacher training, as well as with qualitative and mixed method research design. Together with Lijing SHI, she has innovated the use of eyetracking in a sociocultural framework.