



# **Learner agency in online Task-Based language learning for spoken interaction**

Tesis presentada por

**Janine Knight**

para obtener el grado de Doctor

en Educación y TIC (e-learning)

**Directors:**

**Dr. Elena Barberà**

Universitat Oberta de Catalunya  
(UOC)

and

**Dr. Melinda Dooly**

Universitat Autònoma de Barcelona  
(UAB)

Programa de Doctorado en Educación y TIC

2018

## Thanks

This work could not have been possible without the help of many of you. Firstly, I want to thank my supervisors Elena Barberà (UOC) and Melinda Dooly (UAB) for accompanying me on this path with constant support, encouragement, much patience and respect throughout. Without them I am very sure that my first experience with the research community would not have been as positive and life changing as it has been. Secondly, I thank the *Universitat Internacional de Catalunya* (UIC) for giving me the opportunity of undertaking such a challenge as an employee within their service. I also thank all my brilliant and witty colleagues at the UIC who not only encouraged me throughout but showed unwavering belief that I could complete the work started, even when my own doubt set in. Lastly, I thank my husband, Simon, who made the completion of this work physically possible.

I dedicate the work to my son, Tomàs, who at the time of writing this is 8 years old. This is in the hope that you, through your own life, use your agency wisely, both in your relationships with humans and with computers. I hope that you make choices and act for the common good and attempt to tackle some of the societal and environmental problems that humans face going into the future. However, your purposes are your own.

## **Abstract (English)**

The present study aims to explore the relationship between learner agency, screen-based resources (such as navigational buttons, textual task instructions) and meaning making in synchronous computer mediated communication (SCMC) tasks developed to promote spoken interaction. Using a case-study approach, three tasks designed for language learning (opinion-sharing, role-play and information gap) across two data sets (12 cases) are analysed. Tasks are carried out in an online university in Barcelona with spoken interaction facilitated through an audioconferencing tool. Data was collected over the course of one semester (2015) and analysed, alongside data from a prior study (2012).

The objectives were threefold: to understand how learners' choices and intentional actions pertaining to screen-based resources shape oral turns; to understand how meaning making can be understood from a multimodal perspective, beyond the lingual. The study also aims to contribute to theory on agency in language learning in order to help foster agency in current and future SCMC tasks for optimal language learning gains. A range of data sources and methods are used. Sources include audio recordings of peer-to-peer oral interaction, transcripts, screenshots, course documentation and supplementary information about the technological tool. These sources are explored through data analysis, including content and discourse analysis and Computer Mediated Discourse Analysis (Herring, 2004). In addition, a specific approach is developed that combines emic (learner) and etic (researcher) analytical perspectives that draw on notions from Conversational Analysis (Sacks, Schegloff and Jefferson, 1974) and Multimodal (Inter)actional analysis (Norris, 2004).

Results suggest that types of agency are manifest in tasks. In addition, learners' mediation with screen-based resources are found to shape their oral turns qualitatively and quantitatively. Meaning making involving multiple modes beyond speech (i.e. somatic, text and image) are identified, leading to the implication that agency can be understood as being carried out through human systems (motor, sensory and language) and resources pertaining to the digital system. Agency in SCMC tasks therefore, can be described as 'systems with tool(s)-mediated goal(s)-directed action(s)' which builds on the sociocultural notion of 'tool-mediated goal-directed action' (Zinchenko, 1985). Implications for task design are discussed and recommendations for future research on task-based SCMC from a multimodal perspective are outlined.

**Key words:** Learner agency, screen-based resources, multimodal turn-taking, spoken interaction tasks, Synchronous Computer Mediated Communication (SCMC), semiotic mediation

## Resumen (Spanish)

El presente estudio tiene como objetivo explorar la relación entre el poder de decisión y acción del aprendiz (*learner agency*), recursos basados en la pantalla (como por ejemplo, botones de navegación, instrucciones de tareas textuales) y la creación de significado en tareas de comunicación sincrónica mediada por ordenador (SCMC), desarrolladas para promover la interacción oral. Utilizando un enfoque de estudio de casos, se analizan tres tareas diseñadas para el aprendizaje de idiomas (intercambio de opinión, juego de roles y falta de información en dos conjuntos de datos (12 casos). Las tareas se llevan a cabo en una universidad en línea en Barcelona mediante interacción oral, a través de una herramienta de audioconferencia. Los datos se recopilaron a lo largo de un semestre, en un curso (2015) y se analizaron junto con los datos de un estudio anterior (2012).

Los objetivos eran tres: comprender cómo las elecciones de los alumnos y las acciones intencionales relacionadas con los recursos basados en la pantalla moldean los turnos conversacionales comprender cómo se puede entender la creación del significado desde una perspectiva multimodal, más allá de lo lingual y contribuir a la teoría de *agency* en el aprendizaje de idiomas. La finalidad de este último ha sido ayudar a fomentar *agency* en las tareas actuales y futuras del SCMC y para obtener ventajas óptimas en el aprendizaje de idiomas. Se usa una variedad de fuentes de datos y métodos. Las fuentes incluyen grabaciones de audio de interacción oral punto a punto, transcripciones, capturas de pantalla, documentación del curso, e información suplementaria sobre la herramienta tecnológica. Estas fuentes se exploran a través del análisis de datos, incluido el análisis del contenido y del discurso, así como *Computer Mediated Discourse Analysis* (Herring, 2004). Además, se desarrolla un enfoque específico que combina las perspectivas analíticas emic (alumno) y etic (investigador) que se basan en las nociones del *Conversational Analysis* (Sacks, Schegloff y Jefferson, 1974) y *Multimodal (Inter) actional Analysis* (Norris, 2004).

Los resultados sugieren que los *types of agency* se manifiestan en tareas. Además, la mediación de los alumnos con recursos basados en la pantalla moldean los turnos conversacionales tanto cualitativamente, como cuantitativamente. La creación de significado implica múltiples aspectos más allá del habla es decir, somático, texto e imagen) lo que implica que se puede entender que *agency* se lleva a cabo a través de sistemas humanos (motor, sensorial y de lenguaje) y de recursos pertenecientes al sistema digital. Por lo tanto, *agency* en las tareas del SCMC se puede describir como 'sistemas con herramientas mediados por instrumentos dirigida hacia acciones' que se basa en la noción sociocultural de 'acción dirigida hacia un objetivo y mediada por

instrumentos' (Zinchenko, 1985). Se discuten las implicaciones para el diseño de tareas y se describen recomendaciones para futuras investigaciones basadas en tareas SCMC desde una perspectiva multimodal.

**Palabras clave:** el poder de decisión y acción del aprendiz (*learner agency*), recursos basados en la pantalla, la toma de turnos multimodales, tareas para interacción oral, CMC síncrona (SCMC), mediación semiótica

## Resum (Català)

L'objectiu d'aquest estudi és explorar la relació entre el poder de decisió i d'acció de l'alumne (*learner agency*), els recursos en pantalla (botons de navegació, instruccions escrites per a tasques) i la creació de significat en tasques de comunicació sincrònica mitjançant ordinador (CSMO) orientades a fomentar la interacció oral. El projecte es basa en l'estudi de casos i s'analitzen tres tasques dissenyades per a l'aprenentatge de llengua (intercanvi d'opinions, joc de rols i buits d'informació) en dos conjunts de dades (12 casos). Es tracta d'unes tasques fetes en una universitat en línia situada a Barcelona i per a les quals es va fer servir un sistema d'audioconferència per a facilitar la interacció oral. Les dades es van recollir al llarg d'un semestre (2015) i es van analitzar juntament amb dades recollides en un estudi previ (2012).

L'estudi presenta tres objectius: d'una banda, entendre com les decisions dels alumnes i les accions deliberades pròpies dels recursos en pantalla modelen els torns de paraula; d'altra banda, entendre com la creació de significat es pot concebre des d'una perspectiva multimodal, més enllà de la perspectiva lingüística; en tercer lloc, l'estudi vol ser una contribució a la teoria sobre l'agentivitat en l'aprenentatge de llengües per fomentar l'agentivitat en les tasques CSMO d'avui dia i del futur, a fi i efecte d'aconseguir avançar d'una manera òptima en l'aprenentatge d'una llengua. S'utilitza una sèrie de fonts de dades i de mètodes. Les fonts inclouen gravacions àudio d'interaccions orals entre estudiants, transcripcions, captures de pantalla, documentació de cursos de llengua i informació addicional sobre eines tecnològiques. Les dades s'analitzen a través d'un anàlisi del discurs i de continguts, i d'un anàlisi del discurs mitjançant a través d'un ordinador (Herring, 2004). A més, es crea un enfocament específic que combina les perspectives analítiques èmica (alumne) i ètica (investigador), que se serveixen de l'anàlisi conversacional (Sacks, Schegloff i Jefferson, 1974) i de l'anàlisi (inter)accional multimodal (Norris, 2004).

Els resultats indiquen que en les tasques es manifesten alguns tipus d'agentivitat. A més, el fet que la comunicació entre els alumnes tingui lloc a través de recursos en pantalla modifica els torns de paraula tant a nivell qualitatiu com quantitatiu. També s'ha pogut identificar la creació de significat a través de diversos modes més enllà de la llengua (per exemple, somàtic, text i imatge). L'agentivitat, doncs, es manifesta mitjançant sistemes humans (motor, sensorial i lingüístic) i recursos que formen part del sistema digital. Per tot això, en les tasques CSMO, es pot definir l'agentivitat com aquell «sistema que compta amb accions enfocades a un(s) objectiu(s) i que es desenvolupen mitjançant certa(es) eina(es)», una definició que es basa en la noció sociocultural d'«acció enfocada a un objectiu i desenvolupada mitjançant eines» (Zinchenko, 1985). Finalment, es presenten algunes conseqüències a l'hora de dissenyar tasques i es proposen algunes recomanacions per a futures investigacions en CSMO basada en tasques des d'una perspectiva multimodal.

**Paraules clau:** el poder de decisió i d'acció de l'alumne (*learner agency*), recursos en pantalla, torns multimodals, tasques d'interacció oral, comunicació sincrònica mitjançant ordinador (CSMO), mediació semiòtica

# Index

	Page
<b>Dedication and thanks</b>	
<b>Abstract (English)</b>	
<b>Resumen (Spanish)</b>	
<b>Resum (Catalan)</b>	
<b>Table of contents</b>	
<b>1. Introduction</b>	<b>1</b>
<b>2. Theoretical Framework</b>	<b>2</b>
2.1 Learner agency.....	3
2.2 Multimodality and Mode .....	4
2.3 Semiotic resources and screen-based resources .....	6
2.4 Meaning making in SCMC for spoken interaction .....	8
<b>3. Justification for study</b>	<b>10</b>
<b>4. Objectives and the compendium of papers</b>	<b>11</b>



<b>5. Significance of the study</b>	<b>14</b>
<b>6. Research questions</b>	<b>19</b>
<b>7. Context</b>	<b>20</b>
Table 1. Overview of papers, approaches and methods, foci, tools and research questions .....	24
<b>8. Method</b>	<b>27</b>
8.1. Participants .....	27
8.2. Data .....	27
8.3 Technological Tool.....	28
8.4 Tasks .....	29
<b>9. Approach and methods for analysis</b>	<b>31</b>
Table 2. Data sets and data treatment .....	33
9.1. Coding process for identifying types of agency .....	36
9.2 Methods and approaches in the new focus .....	39
9.3 Changes made to the original focus .....	44
<b>10. Analysis procedure</b>	<b>46</b>
<b>11. Results and discussion</b>	<b>51</b>
Table 3. Summary of main results .....	52

11.1.Results for research question 1 .....	56
11.2.Results for research question 2 .....	57
11.3.Results for research question 3 .....	61
11.4.Results for research question 4 .....	62
Figure. 1	
Higher and Lower-Level Actions in an audioconferencing task.....	64
<b>12. Conclusion to main research question</b>	<b>65</b>
<b>13. How the findings relate to other studies</b>	<b>69</b>
<b>14. Limitations and future research</b>	<b>72</b>
14.1 Limitations .....	72
14.2 Future research .....	72
14.2.1 Agency .....	74
14.2.3 Language learning with a SCMC audioconferencing tool .....	76
14.2.3 Meaning making as verbal and non-verbal in SCMC tasks .....	77
14.2.4 Turns and turn-taking in SCMC .....	79
14.2.5 Mode and task-based learning .....	80
<b>15. Task design considerations</b>	<b>81</b>
<b>16. Bibliography</b>	<b>85</b>
<b>17. Articles in journals or conference proceedings</b>	
17.1 Paper 1, published (first quartile): Knight, J., Barberà, E., & Appel, C. (2017). A framework for learner agency in online spoken interaction tasks. <i>ReCALL</i> , 29(3), 276-293. doi:10.1017/S095834401700009X	
17.2 Paper 2, published (peer reviewed and submitted for inclusion in SCOPUS): Knight, J., & Barberà, E. (2016). The Negotiation of Shared	

and Personal Meaning Making in Spoken Interaction Tasks. In Pixel (Ed.), *Conference proceedings of the 9<sup>th</sup> International conference of ICT for language learning* (p. 248-252). Padova, Italy: libreriauniversitaria. it edizioni.

17.3 Paper 4, manuscript accepted for publication in ReCALL Journal: Knight, J., Dooly, M. & Barberà, E. (2018). Navigating a multimodal ensemble: mediating turns verbally and non-verbally in online interaction.

17.4 Paper 5, manuscript accepted for publication in System Journal: Knight, J., Dooly, M. & Barberà, E. (2018). Multimodal meaning making: navigational acts in an online speaking task.

17.5 Paper 6, manuscript accepted for publication in CALL journal: Knight, J., Dooly, M. & Barberà, E. (2018). But the computer say me the time is up: the shaping of oral turns mediated with and through the screen.

## **Appendix**

### **Appendix 1. additional articles in journals**

Paper 3, published, (second quartile): Knight, J., & Barberà, E. (2017). Learner agency and its effect on spoken interaction time in the target language. *JALT CALL Journal*, 13(3).

Paper 7, in press, (second quartile): Knight, J., & Barberà, E. (2018). Navigational acts and discourse: fostering learner agency in computer-assisted language learning. *Electronic Journal of e-learning*.

### **Appendix 2. codes used for tracking language functions across three task types**

### **Appendix 3. glossary of terms**

---

# **Learner agency in online Task-Based language learning for spoken interaction: how agency manifests and intersects with screen-based resources.**

## **1. Introduction**

Task-based language learning has drawn both researchers' and practitioners' attention since it gained considerable popularity in the 1980s. The use of technology, alongside a task-based approach, continues to be utilised by practitioners as they experiment and adapt the approach, thereby “contributing to the maturation process of both its theoretical conceptualization and practical implementation in foreign and second language education” (Lai and Li, 2011, p.498). According to many authors (Lai and Li, 2011; Thomas and Reinders, 2010; González-Lloret and Ortega, 2014 and Ziegler, 2016), the power of technology has been found to be mutually beneficial for task-based language learning and technology, contributing significantly to learners' language learning processes and outcomes<sup>1</sup>.

Furthermore, communication through a screen has a greater role in modern society than ever before (Liou, 2011), including screens pertaining to computers. Computer-mediated communication (henceforth CMC), can offer learners an ‘ever expanding semiotic budget’ (Blin and Jalkanen, 2014) during task processes. This is due to the evolving and expanding range and complexity of technological tools to facilitate CMC, which allow for example, the use of text chat and video chat in the same tool and can involve a variety of multimedia resources such as audio clips and videos. The range of tools available for synchronous

---

<sup>1</sup> Task process is understood as ‘task-as-process’ (Breen, 1987; 1989) which is concerned with what actually happens in the task. The term ‘task outcome’ is understood as pedagogical task completion (*Skehan*, 1998) so that language is used and learners have completed the task successfully (e.g. spotted all the differences, told a story successfully) (Ellis, 2003). The concepts are taken up in paper 1.

computer mediated communication (henceforth SCMC) or “real-time chat style communication” (Thorne, 2008, p.1), which is the focus of this study, now includes Skype, WebEx, Google Hangout and Face time to name but a few. Consequently, the range of tools and resources on the screen in evolving communicative environments heightens the complex, multimodal opportunities learners can be faced with in current and future task-based scenarios.

Not only can learners be faced with an ‘ever expanding semiotic budget’ (Blin and Jalkanen, 2014), but learners may also be positioned in ways in which they are required to be ‘semiotic initiators and responders’ (Coffin and Donohue, 2014) of resources on the screen. SCMC tasks for spoken interaction not only require that learners communicate and negotiate orally for task completion but they may also provide learners with opportunities to respond to a text or initiate uploading a photo or raise their hand to indicate desire to talk. This scenario conceivably requires that learners make choices and act on these choices in relation to the semiotic resources on the screen, and use them for optimal gains in the target language. Therefore, agency on the part of learners, and the various modes involved in SCMC tasks e.g. speech, image, text, gesture, gaze and somatic (pertaining to the body) are conceivably increasingly central to understanding communication and representation in tasks and the meaning making that can occur ‘beyond the lingual’ (Block, 2013).

## **2. Theoretical Framework**

While the definitions and conceptualisations of key concepts in this study are discussed in more detail in the papers presented in the attached compendium, it is important to highlight some key concepts at the outset. In this section, ‘learner agency’, ‘multimodality’ and ‘mode’, ‘semiotic resource’ and ‘screen-based resource’ and ‘meaning making’ are presented and discussed in relation to task-based, SCMC for spoken interaction. A glossary of terms is also provided in the appendix (Appendix 3).

## 2.1 Learner agency

There is a plethora of definitions regarding learner agency. Martin's (2004) definition of agency as "the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives" (Martin, 2004, p.135) is used in this study, while also acknowledging that agency is "socioculturally mediated" (Ahearn, 2001, p.112). This conceptualisation is used because it allows for an analysis of agency as intentional behaviour during the task process.

Studying learner agency, where learners are understood as social agents, is important for a number of reasons. Not only is learner agency deemed to have inherent value both within and beyond formal learning environments and leads to superior learning (Schwartz and Okita, 2009) but it has also been argued that without learner agency, attainment<sup>2</sup> in second language acquisition (henceforth SLA) cannot be achieved (Pavlenko and Lantolf, 2000). Indeed, as aforementioned, fostering agency is one of the main aims of 'Integrative CALL' (Gruba, 2004) that is based on multimedia and the internet, including hypermedia (Warschauer and Healey, 1998; Warschauer, 2000). Learner agency is also important for developing communication skills and identity work in language learning using SCMC (Kohn and Hoffstaedter, 2017) and should be taken into account when conceptualising digital literacies (Gillen, 2014). In essence, agency is a fundamental construct in understanding learning processes and identities (Miller, 2012).

Additionally, pertaining to tasks facilitated by CMC in particular, CMC can conceivably encourage active participation<sup>3</sup> in task performance and this process can be a means for supporting learner agency in the learning process (Kenning,

---

<sup>2</sup> Understood as 'engagement' in synchronous, collaborative spoken interaction.

<sup>3</sup> While 'active participation' is linked to agency, the study also recognizes that agency can be expressed through deliberate non-participation or non-action (Mercer, 2012) as well as resistance to language learning (Pavlenko and Norton, 2007; Vitanova, 2005).

2010). This is because tasks involving technology, including CMC, imply that learners have to draw not only on their own linguistic and non-linguistic resources but also because they have to draw on their digital skills (González-Lloret and Ortega, 2014). Further connected to the notion of digital skills is the role of multimodal artefacts in the exercising of learner agency. As highlighted by Blin and Jalkanen (2014), tools and resources e.g. wikis, blogs and podcasts can provide students with opportunities to construct and expand the given objects in different, yet converging ways, allowing them to be agents of their own learning.

With regard to how learner agency can be carried out, this study combines concepts from the fields of linguistics and psychology. From linguistics, the notion of language as action or as a way of carrying out 'speech acts' (Austin, 1962; Searle, 1969) is employed. In addition to drawing on the concept that language is a system (Halliday, 1978) and can be used to perform intentional actions, another concept is drawn upon from the field of psychology. This is namely that agency can be carried out through people's sensory and motor systems (Bandura, 1999; Harré & Gillet, 1994). Studies pertaining to learner agency at the level of speech (Novick and Sutton, 1997; Van Lier, 2008; Swain, 2006; Garcia, 2009) and in relation to tasks (Roebuck, 2000; Wang, 1996; Coughlan and Duff, 1994) are presented in paper 1.

## **2.2 Multimodality and Mode**

The concept of Multimodality has been and continues to be articulated and operationalized differently both across different disciplines and research traditions (Jewitt, Bezemer and O'Halloran, 2016). In simple terms, multimodality is an inter-disciplinary approach that understands communication and representation to be more than about language. 'Mode' can be understood as a set of socially and culturally shaped resources for meaning making that have distinct affordances (Kress, 2014) other than language. Modes can include speech, text, image or somatic (pertaining to the body such as touch), moving image and sound, as well as the mode of gaze, gesture and posture in embodied interaction.

Multimodality in SLA is increasingly becoming an established research focus, involving a vast array of computer-assisted language learning scenarios, including CMC (see ReCALL, special edition, September, 2016). Many researchers highlight the need to take into account multimodality when it comes to task design (Berglund, 2009; Canto, de Graff and Jauregi, 2014; Hampel, 2006, 2010; Hauck, 2010 and Stockwell, 2010).

Both researchers within the fields of SLA and social semiotics have highlighted a need for a methodological 'turn' in analysing learning scenarios to shift from a purely language focus to analysing speech alongside other modes. This is in order to understand interactions more holistically (Kress and van Leeuwen, 2001; Lamy, 2006; Hampel and Hauck, 2006; Flewitt, 2008; Flewitt, Hampel, Hauck and Lancaster, 2017).

A brief overview of how the focus on multimodality and mode have been approached in SCMC for SLA is useful to position the current study. Ziegler (2016), noted that initial explorations of multimodal learning contexts and tasks focused on the suitability, design, and implementation of tasks, with the finding that environments integrating more than one type of technology can and do provide learners with opportunities to receive and provide feedback, to focus on form, and for second language (henceforth L2) development (Hampel, 2006; Wang, 2006). As research on multimodality has evolved, SCMC studies including those with a focus on spoken interaction, have approached multimodality through the notion of 'mode-as-communication channel' or 'modalities' (e.g. text chat, videoconferencing, voice chat/audioconferencing) which when combined together, through same or different tool use, can also contribute to bi-modes or tri-modes for communication. Studies on SCMC for spoken interaction for language learning, some task-based, that have taken this approach include Satar and Ozdener (2008); Jepson (2005); Sauro (2009); Jung (2016); Bueno-Alastuey (2013); Sykes (2005) and Yanguas (2010).

Other studies on multimodality and SCMC for spoken interaction have highlighted specific resources on the screen and/or their modes, such as an image or text (involving the modes of image and text), yes/no buttons (involving the modes of



image, text and somatic mode), a moving image (involving the mode of image) etc.

The papers presented in this compendium, focus on multimodality from this second approach and the studies pertaining to this focus are subsequently presented. More specifically, the papers in this study focus on 'screen-based resources' that is a new term used by Balaman and Sert (2017a; 2017b) and which they use to refer to resources such as photos, videos or texts that pertain to the interface.

### **2.3 Semiotic resources and screen-based resources**

'Semiotic resources' in sociocultural theory include print materials, the physical environment, gestures and classroom discourse: in essence any resources that mediate the developmental process of language learning (Wertsch 1991; 1998). From this perspective, the goal of task completion can be achieved through mediational means using any number and/or combination of semiotic resources. Semiotic resources, including screen-based resources, are made up of modes: spoken language is both aural (the mode of sound) and oral (the mode of speech). A text on the screen for example, can pertain to the mode of text and image (due to the layout and positioning on the screen). External semiotic resources correspond with the notion of 'input' as task features (see Candlin, 1987 and Robinson, 2011 for overview). The relevance of 'input' has been recognized by both cognitivist researchers on task-based learning (e.g. Robinson, 2011; Skehan, 1996) as well as researchers that have taken a sociocultural perspective to SLA (e.g. van Lier, 2000). Input is "the written, visual, or aural information that learners performing a task work on to achieve the goal of the task" (Robinson, 2011, p.7). 'Input' can include task instructions and the input which learners are required to process and use (Ellis, 2000) such as pictures, a map or written text (Ellis, 2003)<sup>4</sup>. In order to acknowledge both the

---

<sup>4</sup> In online tasks, task instructions, as a case in point, can pertain to technological aspects (e.g. textual or auditory instructions for learners to start the task by clicking something) or pedagogical aspects (e.g. textual instructions pertaining to learner roles).

interactional and mediational aspects of CMC, the term 'screen-based resources' rather than 'task features' or 'inputs' is used in this study.

While SCMC tasks facilitated by the use of audioconferencing tools, which is the focus of this study, can be considered "voice but no image" (Yamada, 2009) because learners do not see their peer synchronously and the channel is monomodal (audio only), Collentine (2009) proposes that learners have the possibility of having a rich, multimodal experience. This is because the information represented to learners can involve spoken language from their peers but also involve being faced with various screen-based modes and resources. Whereas the embodied modes (pertaining to the human body) of their peer or teacher such as gaze, gesture and posture are not present, these can be 'replaced' by non-embodied modes (pertaining to a screen rather than humans) while learners converse.

A number of SCMC studies for spoken interaction highlight the screen-based modes and resources both pertaining to audioconferencing, videoconferencing and audio graphic conferencing tools. Collentine (2009) identified the modes (and screen-based resources) of text and image in an online task-based audioconferencing event. Lamy's (2006) study on turn-taking and face-saving using an audio-graphic tool identified natural language (spoken and written) as well as visual resources (icons, images, colours and shapes) and found that a turn sequence could be performed transmodally i.e. across text and oral speech. Vetter and Chanier's (2006) study on how language learners used multimodal tools to make spoken interactions highlighted text, speech, graphics for communication and highlighted the interplay of modes (text and spoken language). Kenning (2010), also highlighted the various features of two different audioconferencing tools and noted their heterogeneity with regard to their 'functionalities' such as colour, a particular search facility, a shared tool and a voice activated command. Hampel and Stickler (2012), in a videoconferencing event, identified communication modes as linguistic (spoken and written) alongside visual such as icons (vote buttons yes/no/? emoticons), still and moving images, display/scrolling of text and gestures. Lamy (2012) identified the clicking of a button as a means to say goodbye. More recently Balaman and Sert's

(2017a) study on two different task types in two different settings (face-to-face and online using audioconferencing) highlighted the video clips on the screen whereby learners could type answers, click on answer buttons and receive correct answers whilst conversing with their partner. Whereas in the face-to-face task learners orientated to topic maintenance, in online tasks orientation was to the interface and on screen activities. In another task-based study, Balaman and Sert (2017b) confirmed how participants coordinated their (oral) interactions with their orientations to the screen-based resources on the task interface.

## **2.4 Meaning making in SCMC for spoken interaction**

Tasks, by definition, make meaning primary but require language to enable meanings to be communicated (Skehan 1992, 1994) whereby learners should be mainly concerned with processing the semantic and pragmatic meaning of utterances (Ellis, 2003). In sociocultural theory, it is through “collaborative dialogue” (Swain, 2000, p. 112) that meaning is constructed through the mode of speech. Meaning is not necessarily something given and accepted, but rather learners (orally) “negotiate the message of the input” (Swain, 2000, p. 98), in the process of dialogue.

Oral meaning making in task-based SCMC has been a central focus of study from many different perspectives, research foci and methodological approaches. Three strands of research relating to oral meaning making in SCMC for spoken interaction that emerged as pertinent to the data and research focus in this study include: 1) Negotiation of Meaning that is concerned with quantity of meaning making, measured using oral turns; 2) discourse that relates to the content and purposes of oral utterances - often analysed with a focus on functions or speech acts and 3) oral turn-taking and turns<sup>5</sup> that can be concerned with how knowledge, or meaning, is co-constructed.

The first, ‘Negotiation for Meaning’ (henceforth NfM), has been an important focus in cognitivist approaches to SCMC (see Sauro, 2011) and task-based research

---

<sup>5</sup> ‘Turn-taking’ is a type of organisation in conversation whereas ‘turns’ pertains to individual oral turns within the organisation.

(Ziegler, 2016). NfM focuses on learners' oral turns for clarification requests and comprehension checks, among others, in order to overcome comprehension difficulties (Varonis and Gass, 1985) and has been researched extensively with regard to SCMC tasks. The total number of negotiated turns in a negotiation of meaning sequence are measured, in order to ascertain how various task types in SCMC, for example, can stimulate a greater amount of talk in the target language (e.g. Jung, 2016). While this study does not adopt a cognitivist approach, the study is concerned with the quantity of oral interaction in the target language as a task outcome because of the purported language gains made through interaction (de la Colina and Mayo, 2007; Gass and Mackey, 2006). Meaning making from this perspective i.e. quantitative measure of spoken interaction in the target language, can be important for task design for SCMC for spoken interaction, because achieving extended oral interaction in the target language may be a main design objective for spoken interaction tasks.

Secondly, speech acts and discourse functions have also been a focus pertaining to oral meaning making in SCMC. Collentine's (2009) study of an SCMC task for spoken interaction identified interpersonal and sociocultural behaviours (e.g. joking, off-task discussions) as part of task talk. In a synchronous, non-task based CMC language learning scenario, Cunningham, Fägersten and Holmsten (2010) and Fägersten, Holmsten and Cunningham (2010) identified meta-modal discourse pertaining to the screen-based resources of a seminar tool involving video, whiteboard facility and text chat. Meta-modal communication occurred in relation to tool use (the pointer function, text box) (Cunningham, Fägersten and Holmsten, 2010) and to initiate repair work and facilitate the progression of communication (Fägersten, Holmsten and Cunningham, 2010).

Thirdly, meaning making, understood as knowledge co-construction, has also been studied by focusing on oral turn-taking in SCMC (see Gonzalez-Lloret, 2011; Jenks, 2014). A number of CMC studies, both focused on language learning (including task-based) and non-language learning scenarios for spoken interaction, have highlighted how clicking various screen-based resources can form part of the accomplished interaction and therefore meaning making. Lamy (2012) found that using an audio graphic conferencing tool, a user may close a

conversation by typing 'bye for now', by clicking a specific button, or withdrawing orally (Lamy, 2012) and proposes that clicking a resource that means 'temporarily away' is likely to influence the direction of the conversation differently compared to disconnecting altogether. Liddicoat (2010) found that in an opening videoconferencing exchange the securement of a non-present co-participant for the conversation was achieved by the initiator through the technology: a message via the computer (screen). According to Liddicoat (2010), this sequence resembled summons-answers sequences. Furthermore, whilst learners were establishing a channel, learners orientated to the technological interface in their talk, rather than to personal interaction with their peer before moving to the first topic.

In the field of social semiotics, meaning making is understood to occur in a variety of ways, not only with language. Hampel and Hauck (2006) note the importance of the role of different modes for meaning making in CMC, noting the growing importance of the visual mode with a shift from the book to the screen. They also highlight a theoretical framework of multimodal meaning making, building on the work of Kress (2000) and Stein (2004), whereby people are seen as agents who are making meaning and producing texts and constantly remaking the representational resources in the process.

### **3. Justification for study**

Looking at SCMC tasks from a purely lingual perspective limits our understanding of multimodal communication in SCMC tasks. Research is needed to understand the impact of multimodal communication in online classrooms (Hampel and Stickler, 2012; Dooly and Hauck, 2012) and mode in task-based language classrooms in general (Gilabert, Manchón and Vasylets, 2016).

Furthermore, given that fostering learner agency can be an aim of task design, including those involving SCMC tools, understanding how agency manifests in the process of task completion is important in order to know how to foster it. In addition, how agency and screen-based resources shape oral interaction, both qualitatively and quantitatively, would offer insight into multimodal communication

in SCMC tasks and also offer insight for task design that seeks to stimulate extended, synchronous, oral interaction in the target language.

Related to the role of modes in the process of meaning making, is the agency of the meaning makers as social agents (Bezemer and Kress, 2016). Therefore, in order to understand how the screen-based resources can contribute to meaning making beyond the lingual, the relationship between learners as social agents and screen-based modes during task process needs further exploration. Because semiotic resources, each offer distinct potentialities and limitations (Jewitt, Bezemer and O'Halloran, 2016), the potentialities for meaning making of the screen-based resources are also under focus.

Finally, and important to the context of the study, is the shift in the research of task-based learning and technology from purely cognitive studies to those with a sociocultural focus that recognizes that many new technologies are essentially social in nature (e.g. Hellermann, Thorne and Fodor, 2017; Dooly, 2015; 2016; Block, 2003). Due to this shift, tasks are more recently being studied in relation to other learning outcomes such as intercultural competence and digital literacies (Gonzalez-Lloret and Ortega, 2014). Adding a multimodal lens to a sociocultural perspective on tasks also brings a focus on the meaning making potential of all screen-based modes, not just speech to the fore, thus raising the question as to what meaning making is occurring 'beyond the lingual' (Block, 2013).

#### **4. Objectives and the compendium of papers**

The study, presented as a collection of papers, looks at different aspects involved in three Computer-Mediated Communication (CMC) tasks that were designed to specifically elicit spoken interaction. The interaction was facilitated through an audioconferencing tool, along with other online resources (e.g. documents, pop-up instructions, etc.). This compendium of papers aims to further researchers' and practitioners' understanding of the relationship between agency and screen-based resources in this type of interaction and the meaning making that emerges from this process. More specifically, it aims to examine the relationship between learner choices and how learners act on these choices in relation to screen-based

resources from their peer and on the screen and how learner oral interaction is shaped both qualitatively and quantitatively in the process of mediated action for task completion<sup>6</sup>.

Overall, the study has a number of research foci. The research seeks to understand how agency is manifested in a CMC task-based scenario in order to inform online task design. Perspectives on agency are explored in paper 1 and paper 2 with the focus on learner agency and the mode of speech. Papers 3, 4, 5 and 7 relate to agency and the somatic mode (pertaining to the body). Papers 2, 4 and 6 also focus on the modes of image and text.

The study also aims to further understand how learners' intentional choices and actions based on these choices (understood as agency, thus overlapping with the previous aim mentioned above). This is explored by taking into consideration how screen-based resources (which go beyond speech) may shape learners' oral turns in CMC tasks for spoken interaction, particularly the somatic mode pertaining to navigation. Agency and the role of the somatic mode, (pertaining to the human body), in shaping talk, is addressed qualitatively in papers 1, 2 and 6. Agency and the role of the somatic mode in shaping talk are addressed through a quantitative lens in papers 1, 3 and 5.

Because the thesis aims to inform future online task design that seeks to foster agency as well as support language learning gains, the study intends to gain insight into learners' mediation process with different screen-based resources, including navigational resources. Specifically, this is addressed in paper 4, which outlines how mediation with screen-based resources, supports the carrying out of agency. This can give greater insight into mediation with screen-based resources by taking into account the complexity of screen-based modes, the expanding semiotic budget related to the screen, and the materiality of the features of tasks.

---

<sup>6</sup> Submitted manuscripts (papers 4, 5, 6) and the paper accepted for publication (paper 7) require final checks, revisions and minor amendments before publication. Therefore, for future research purposes, the final published versions should be referred to or cited, not the manuscript versions in the compendium.

Because there is also a need to understand the complexity of all the modes involved in meaning creation (Calvo Ferrer, Melchor-Couto and Jauregi, 2016), including during tasks, the study also aims to understand how meaning making in task-based CMC for spoken interaction is constructed beyond the lingual. This means not only exploring meaning making through speech in papers 2 and 6, but also non-verbal meaning making which is taken up in paper 5. By focusing on how learners use all the screen-based resources for meaning making in the action of task completion, including modes such as image and text, an objective is to give insight into future task design and other potential learning gains, such as personalization of learning, that may be obtained through their use.

The study also contributes to an understanding of how screen-based modes impact online tasks, specifically in relation to the shaping of spoken interaction, both qualitatively and quantitatively, as well as deepen the understanding of the relationship between mode and meaning making and agency.

Parallel to these aims, the study seeks to contribute to multimodal methodological approaches for analysing SCMC for audioconferencing. Although advances in eye tracking technologies for SCMC (whereby learners meaning making with the screen-based resources can be explored) have been made (Stickler and Shi, 2017) this method would be a particular challenge given the geographical dispersion of online learners and even so, may not offer insight on covert learner behaviours that the conceptualisation of learner agency used in this study would need to encompass. Methodological approaches are still at an exploratory stage (Rossolatos, 2015) and remain a challenge (Herring, 2015) despite a few key studies (see Flewitt, Hampel, Hauck and Lancaster, 2017; Lamy, 2012 and ReCALL Special Issue, 2016).

Finally, this study is timely given that learner agency is under-theorized, there is a lack of clarity in defining it, difficulty exists in establishing sound analytic research procedures and operationalizing it remains a challenge (Miller, 2012). By seeking to understand learner agency as a “contemporary phenomenon in depth and within its real-life context” (Yin, 2009, p.18), this study can contribute



significantly to current and future research on agency and CMC tasks as well as inform practitioners regarding online task design that seeks to foster learners' agency and support optimal gains in language learning.

## **5. Significance of the study**

This study is significant and makes a number of original contributions to the field of SCMC, task-based learning for spoken interaction, building on the work of others. The contributions correspond to 5 areas: 1) the importance of non-verbal modes in SCMC for spoken interaction; 2) understanding meaning making, including and beyond the lingual; 3) agency in SCMC for spoken interaction; 4) insights for task and tool design and 5) contributions to sociocultural theory and the study of mediation and screen-based resources.

### **The importance of non-verbal modes in SCMC for spoken interaction**

Firstly, the study highlights the importance of non-verbal modes and screen-based resources in SCMC tasks. In particular, the importance of the somatic mode through learners' use with navigational resources with respect to how it shapes learners' oral interactions qualitatively and quantitatively. A key contribution is the finding that some screen-based resources, pertaining to non-verbal modes were also identified as turn-takers and discourse participants.

### **Understanding meaning making including and beyond the lingual**

Secondly, the study contributes to identifying meaning making in SCMC task-based learning as being carried out through oral and non-verbal means through the mediation of screen-based resources, alongside spoken language. Screen-based resources/modes were also identified as shaping how and what was negotiated, offering new opportunities for meaning making, including and going 'beyond the lingual' (Block, 2013). The findings contribute to understanding the role of the screen in the construction and shaping of oral and non-verbal turns.

In relation to speech as mode, different foci for talk were identified beyond the pedagogical task, which pertained to smaller goals in the task such as learners organizing themselves or negotiating navigation. This offers insight into research on Negotiation for Meaning from a cognitivist perspective because oral turns can pertain to technological aspects of the task not only the pedagogical aspects, the latter being the basis of research on NfM sequences and the comparison of task types.

Whereas SCMC for SLA has predominantly focused on oral or written turns and turn-taking, the study identifies 'semiotic initiations and responses' (Coffin and Donohue, 2014) as to how multimodal turn-taking between human-human and human-screen-based resources is carried out. Whereas multimodal turn-taking with embodied modes (e.g. gaze, gesture) has been explored in non CMC scenarios (Mondada, 2007; 2013), this study offers a new contribution by identifying multimodal turn-taking with embodied (speech) and non-embodied screen-based modes. In addition to Mondada's (2013) call for 'a visual turn' in the field of linguistics, there may also be a case for 'a somatic turn'.

The study adds to the growing body of research that highlight how the coordination of interaction with orientations to the online task interface shapes epistemics (knowledge construction) in online interaction (Balaman and Sert, 2017a; Balaman and Sert, 2017b and Musk, in press). While these studies contribute to a multi-layered view of knowledge co-construction (Goodwin, 2013), or meaning making, with the multitude of epistemic, semiotic, interactional, and screen-based resources that the participants simultaneously draw on, this study contributes something new. It identifies how the screen-based resources are not only orientated towards by learners, but that some screen-based resources can act as 'others' (Raudaskoski, 1999): initiate responses from learners, form part of learners' oral turns, and/or act as 'agents' in the mediation process, whereby turn-taking can be understood as being shared between humans and computers. Non-verbal meaning making can be considered to be on the whole "constructed one-sidedly, rather than co-constructed, making the human participant solely responsible for the emerging meaning" (Raudaskoski, 1999, p. 22-23). However,

the computer as a sense maker was also identified in the way that it can 'respond' by taking learners to another page/place through navigational means.

The study brings meaning making in SCMC tasks as a learner experience to the fore. Learners' physical negotiation of the purpose of tool, through different navigational trajectories, meant that learners could 'negotiate' how they experienced the tool (as a form of meaning making) and their roles pertaining to it, in as far as how they positioned themselves as tool users or tool managers. In doing so, they could conceivably construct their identity with respect to who they were, in relation to the tool, and the personal/social meanings they could draw from that process.

New opportunities for oral meaning making were also identified pertaining to navigational resources as a topic of talk, beyond the pedagogical task, as well as a photo and open question (textual) that facilitated personal and (orally) shared meaning making. Meaning making was identified as being a multisensory process, involving sensory (image mode) and motor (somatic mode) systems, as well as the spoken language system (speech mode). Furthermore, the meaning making process could be considered at times, a multi-party encounter between learners and the various visual/textual/navigational and textual/visual screen-based resources.

### **Agency in SCMC for spoken interaction**

Thirdly, the study advances theory on learner agency in online SCMC for spoken interaction, showing how human systems of (oral) language together with the sensory and motor systems were used to negotiate task process and shape task outcome. The study also underscores the intentionality exhibited by some screen-based resources as they 'initiate' in order to obtain a learner response, particularly the intentionality of navigational resources that can be considered as 'others' (Raudaskoski, 1999), acting as and/or being orientated towards by learners as active agents.

This leads to a systems-based understanding of agency whereby actions are shared between systems and screen-based resources and whereby intentionality is negotiated (somatically or visually) with other parts of the (computer) system. Agency can therefore be described as being carried out as 'systems with tool(s)-mediated goal(s)-directed action(s)', building on the sociocultural notion of 'tool-mediated goal-directed action' (Zinchenko, 1985).

Agency was demonstrated through learners' use of language-as-action (speech), through the use and non-use of pedagogical task instructions and prompts (textual/visual) as well as the repurposing of navigational resources and technological tool (somatic) which contributes to a specific understanding of the relationship between agency and screen-based resources (and modes).

The study contributes to the theoretical and methodological challenge of studying agency in SCMC tasks: advancing theory, offering a research procedure and a way to operationalize it.

### **Insights for tool and task design**

Fourthly, the study offers insight for task and tool design by underscoring the importance of learners' purposes for a tool; how learners actually use it versus task design (that also may be different from one use to the next). Furthermore, the study highlights how learners position themselves (e.g. as tool users or tool managers) in relation to tool use and what tool and task designers design learners to *be* in relation to technological tools. Learners' 'undesirable' navigational behaviour in relation to the tool can be understood as 'legitimate peripheral participation' (Lave and Wenger, 1991) as a way of negotiating the rules of tool

use in the community (online) of practice. Learner agency with respect to digital literacy<sup>7</sup> and ‘cultures-of-use’<sup>8</sup> (Thorne, 2016) are highlighted.

The study also brings to the fore, intentional navigational acts as a form of agency, supporting Gillen’s (2014) proposal that agency is central in developing digital literacy. However, in online SCMC tasks this may be simultaneously problematic for task and tool designers that design a tool/task for a specific purpose.

### **Contributions to sociocultural theory, mediation and screen-based resources**

Finally, the study contributes to sociocultural theory in relation to mediation, with the notion of ‘tools within tools’. This notion proposes that screen-based resources are the tools ‘within’ the technological tool. Building on the work of Norris (2004), it proposes that mediation with screen-based resources are ‘lower-level actions’ (Norris, 2004) that constitute the ‘higher-level action’ (Norris, 2004) of mediation with the audioconferencing tool. Also with respect to mediation, the study echoes Kurek and Hauck (2014), who propose that technology enhanced language learning activities, including tasks, are “mediated twice: by the technology used and by the second language (L2)” (Kurek and Hauck, 2014, p. 122). This study deepens and extends this notion by highlighting the dual purpose that learners can also have during SCMC tasks connected to these two lines of mediational means.

---

<sup>7</sup> While there is a range of definitions for digital literacy, the term is used to refer to the use of digital tools and resources to foster the autonomous development of the basic linguistic and intercultural skills required to engage in interactions. (European Centre for Modern Languages, 2017)

<sup>8</sup> Historically shaped practices with particular purposes, associations and values that accrue to a tool from its everyday use (Thorne, 2016). Thorne describes various mismatches between designed and preferred tool use amongst students, different purposes and groups of people.

## 6. Research Questions

The following research questions emerged from the development of paper 1 (section 12) on how learner agency manifested in a synchronous online spoken interaction event, which resulted in 'A Framework for learner agency in spoken interaction tasks' by Knight, Barberà and Appel (2017) (paper 1). We subsequently expand on the key findings from paper 1 in the forthcoming 'context' section, which explains what led to this study. Following on from these findings, and to respond to the research gap regarding a lack of research on mode in task-based online language learning classrooms (including SCMC), the main question that emerged for this study was:

How does learner agency intersect with screen-based resources in SCMC tasks for spoken interaction and what can we learn about agency and meaning making in this relationship?

The sub-questions and papers that address the questions are as follows:

1. What are the foci of learners' goal-directed mediated actions in spoken interaction tasks and how do learners negotiate meaning of the content? (paper 2 in section 12)
2. How does the mediated action with different screen-based resources shape learners' oral turns?
  - Shape learners' oral turns quantitatively (measured as time in the target language), (paper 1 and paper 5 in section 12 and paper 3 in the appendix)
  - Shape learners' oral turns qualitatively (paper 1 and paper 6 in section 12 and paper 3 in the appendix)
3. How do learners carry out the goal-directed mediated action of task completion *through* and/or *with* the available multimodal screen-based resources? (paper 4 in section 12 and paper 7 in the appendix)

4. How is meaning being made with non-verbal modes in online tasks? (paper 2, paper 4, paper 5 in section 12 and paper 7 in the appendix)

## **7. Context**

As aforementioned, an exploratory study on learner agency by Knight, Barberà and Appel (2017) (paper 1) was carried out based on data originally derived from a study by Appel, Robbins, Moré and Mullen (2012) that used a quantitative analysis to explore how different navigation features of a Tandem tool influenced learning strategies and target language use in tasks. The study on learner agency, which took a qualitative approach, used the data from Appel, Robbins, Moré and Mullen's (2012) study (audio recordings of oral peer-to-peer interaction) that were purposively sampled on the basis of dyads who demonstrated 'good' performance or carried out tasks as designed and those that approached the task differently. In Knight, Barberà and Appel's study (2017) (paper 1) the interaction of several dyads were used as case studies and an analysis of the oral interaction confirmed that some dyads had looked at the answers to a task before recording and had planned to some degree, their oral turns. This event became a focus for further exploring and developing theory about learner agency in online tasks that took into account the possibility of learners using their agency in ways that appeared contradictory to their goal of learning a language. In addition, different screen-based resources became part of learner talk or absent from the talk, which warranted further investigation. Because learners were clicking navigational resources to navigate and control the tool, the physical and non-verbal aspects of the task were highlighted i.e. the use of the somatic mode (pertaining to the human body). Some learners also used the photo in the task to talk about themselves orally to their peer in relation to an image of London/ a holiday destination (what they had interpreted the topic to be about).

From these findings, our study on learner agency evolved. Whilst at the beginning a sociocultural perspective to language learning was adopted, including the notion of mediation with tools beyond spoken language as a tool, the focus was still on analysing learner talk as the sole mode and the meanings emerging from that focus. This was the focus in paper 2 (Knight and Barberà, 2016). However,

based on the results from Knight and Barberà's study (2016) (paper 2), it became apparent that in order to fully understand meaning making and agency, including how meaning was negotiated, it could not be done by purely focusing on learners' spoken language. It became apparent that, in addition to spoken language being used as a resource in the task, learners were using other screen-based resources (e.g. navigational resources, photo as image) that pertained to non-speech related modes. Furthermore, the importance of the non-speech related modes was underscored by the fact that some learners used the somatic mode to reconfigure the task design, and this in turn affected their oral interaction (speech as mode).

The meaning making emerging from learners' mediation process was also part of this focus because the personal and shared meanings that learners were making across the three tasks (information gap, opinion-sharing and role-play) appeared intrinsically linked to how learners were using (mediating with) the screen-based resources and the purposes for which they were being used. From this conclusion, the study turned from coding transcripts from learners' audio recordings (to see if types of agency appeared in another task type using content analysis and discourse analysis), to analysing the two data sets multimodally. This meant devising some specific methods for analysis that had not been previously used before in task-based language learning studies nor in the field of multimodal research for SLA, as far that the main author was aware.

Because some dyads across the two data sets explicitly mentioned the navigational resources as a focus of talk, and it was the means whereby learners were controlling the technological tool, the screen-based navigational resources as mediational tools became a particular focus in how the study evolved.

The following description of the methods and procedure undertaken in the study follow on from the results from Knight and Barberà (2016) (paper 2) and so starts with the methods and procedure relating to the data collection and analysis from a multimodal perspective. The data collected and analysed was expanded beyond the learners' audio recordings and transcripts so that learner talk was not



the sole focus. In addition, the interface screen and specific screen-based resources become a main data source and foci for analysis.

The procedures subsequently described, follow the 'multimodal turn' that took place in the study. Whereas the dyads from the second data set (role-play) initially served the purpose of facilitating a cross-case analysis for types of agency across task types, the second data set later became the basis for a new exploratory focus in its own right. This new focus was on learners' choices, and actions based on those choices (agency), in relation to different screen-based resources (made up of modes). There was a need to further explore how the resources were used as tools in the goal of task completion, and how learners were making meaning from this process. In short, the focus taken was on the intersection between agency, screen-based resources and meaning making.

After focusing on the second data set (role-play) to develop the methods and procedure for a multimodal analysis, the first data set was revisited, as a purposive sample (all be it pre-selected as a sample) and which facilitated a cross-case analysis. A visual summary is provided in Table 1 that is an overview of the approaches and methods used for each paper, how they fit together, the tools used and research questions that pertain to each other. It highlights how the original approach and methods changed in the process of studying the phenomenon, taking a 'multimodal turn' in terms of how the phenomenon was analysed and understood. The table also shows that by paper 4, work had also commenced in order to devise specific ways for studying agency (and mediational means) through a multimodal lens for the specific CMC scenario. While the results and foci of paper 3 and paper 7 contribute towards answering the research questions, they are considered as 'additional' in order to simplify and not detract from the main lines of research and findings of the study that emerged.



**Table 1. Overview of papers: approaches, methods, foci, tools and how they pertain to the research questions**

	<b>Approach and methods</b>	<b>Focus</b>	<b>Analytical Tools</b>	<b>Data set(s) and main source(s) for analysis</b>	<b>Research questions</b>
<p><b>Paper 1</b></p> <p><i>A Framework for learner agency in online spoken interaction tasks</i></p> <p>Knight, Barberà, and Appel, C. (2017) (published)</p>	<ul style="list-style-type: none"> <li>• (Exploratory) case study approach</li> <li>• Sociocultural discourse analysis</li> <li>• Etic perspective</li> </ul>	Agency and speech (learners)	Analytical Framework of Peer Group Interaction (Kumpulainen and Muntanen, 1999)	Audio recordings/transcripts One data set (Information gap task followed by opinion-sharing)	RQ2 (shape talk quantitatively) RQ2 (shape talk qualitatively)
<p><b>Paper 2</b></p> <p><i>The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks</i></p> <p>Knight and Barberà, (2016) (published)</p>	<ul style="list-style-type: none"> <li>• Case study approach</li> <li>• Content analysis- coding</li> <li>• Etic perspective</li> </ul>	Agency and speech (learners)	Language functions taken from the Analytical Framework of Peer Group Interaction (Kumpulainen and Wray, 2002)	Audio recordings/transcripts Two data sets (set 1. Information gap and opinion-sharing task plus set 2. Role-play task)	RQ1 RQ4
<p><b>Paper 4</b></p> <p><i>Navigating a multimodal ensemble: mediating turns verbally and non-verbally in</i></p>	<ul style="list-style-type: none"> <li>• (Exploratory) case study approach</li> <li>• Etic and Emic perspective</li> <li>• Multimodal</li> <li>• Tracking using learner speech (relevance)</li> <li>• Task simulation</li> <li>• Reconstruction</li> </ul>	Mediated action 1. speech and physical acts in the form of turn-taking with screen-based resources (learners)	Notions from Multimodal (inter)actional analysis (Norris, 2004) Notions from Conversational analysis (relevance and orientation) (Sacks, Schegloff and Jefferson, 1974) Discourse category of 'turn' applied to physical turns with a screen (Benson,	Audio recordings One data set (set 2. Role-play task) Screenshots from researcher task simulation	RQ3 RQ4

<i>online interaction</i>  Knight, Dooly and Barberà, (submitted manuscript)	----- <ul style="list-style-type: none"> <li>• Labelling of screen-based resources</li> </ul>	----- 2. the semiotic resources on the screen	2015) then applied to navigational clicks as turns ----- Labels for screen-based resources based on work by Lamy (2006) and Herring (2015)	----- Screenshots from researcher task simulation	
<b>Paper 5</b>  <i>Multimodal meaning making: navigational acts in an online speaking task</i>  Knight, Dooly and Barberà, (submitted manuscript)	<ul style="list-style-type: none"> <li>• Case study approach</li> <li>• Etic and Emic perspective</li> <li>• Multimodal</li> <li>• Tracking through learner speech (relevance)</li> <li>• Task simulation</li> <li>• Reconstruction</li> </ul>	Mediated action 1. physical acts through navigation (learners)	Computer Mediated Discourse Analytical perspective (CMDA) (Herring, 2004)	Audio recordings Screenshots from researcher task simulation Tandem tool logs  Two data sets (set 2. Role-play task and set 1. information gap and opinion-sharing task)	RQ2 (shape talk quantitatively) RQ4
<b>Paper 6</b>  <i>Multimodal meaning making: navigational acts in an online speaking task</i>  Knight, Dooly and Barberà, (submitted manuscript)	<ul style="list-style-type: none"> <li>• Case study approach</li> <li>• Emic and etic perspective</li> <li>• Multimodal</li> <li>• Comparison of transcripts and screenshots</li> </ul>	Mediated action 1. speech and physical acts through navigation (learners)  2. the semiotic resources on the screen	Discourse category of 'turn' applied to learners' physical clicks (Benson, 2015)  Notions from Conversational analysis (orientation) (Sacks, Schegloff and Jefferson, 1974)	Transcripts from audio recordings Screenshots from researcher task simulation	RQ2 (shape talk qualitatively)

**ADDITIONAL PAPERS**

<p><b>Paper 3</b></p> <p><i>Learner agency and its effect on spoken interaction time in the target language</i></p> <p>Knight, and Barberà, E. (2017). (published)</p>	<ul style="list-style-type: none"> <li>• Case study approach</li> <li>• Etic perspective</li> </ul>	<p>Agency in relation to technological tools, time mode and language use</p>	<p>Analytical Framework of Peer Group Interaction (Kumpulainen and Muntanen, 1999)</p>	<p>Audio recordings/transcripts</p> <p>One data set (set 1.Information gap task followed by opinion-sharing)</p>	<p>RQ2 (shape talk quantitatively) RQ2 (shape talk qualitatively)</p>
<p><b>Paper 7</b></p> <p><i>Navigational acts and discourse; fostering learner agency in computer-assisted language learning</i></p> <p>Knight and Barberà (in press)</p>	<ul style="list-style-type: none"> <li>• Systematic literature review</li> <li>• Progressive focusing (Stake, 1981)</li> </ul>	<p>Agency, navigation and 'speech' (learners and computers)</p>	<p>Model by Sinkovics and Alfoldi (2012)</p>	<p>CALL scenarios selected from papers in the literature review</p>	<p>RQ3 RQ4</p>

## 8. Method

### 8.1 Participants

Participants were adult learners enrolled on an English as a Foreign Language class as part of their degree programme at the *Universitat Oberta de Catalunya*, a 100% online university based in Barcelona.

There were two data sets: eight participants (four dyads) from the first data set (carrying out an information gap and opinion-sharing task) who had a global level of B2.1 on the Common European Framework of Reference for Languages (CEFRL). Five were female and three were male. The second data set was made up of eight participants (four dyads), (carrying out a role-play task), and had a level B1.1 on the CEFRL. Two participants were male and six were female. The second data set was added to iteratively (making 6 cases) when the focus of analysis turned to learners mediating differently with some screen-based textual resources than what was intended in the task design. In addition, a closer analysis of case 6 added to what was learnt from analysing and choosing the original four cases. The participants were non-native adults, aged between 26-55<sup>9</sup> years old and were considered bilingual, sharing Catalan and Spanish.

### 8.2 Data

The first data set (from Appel, Robbins, Moré, and Mullen, 2012) consisted of recorded oral conversations between four dyads, approximately 23 minutes in total. Audio recordings were made by learners themselves and captured using a plug-in for Skype, a free video and audio conferencing tool that worked with the Tandem tool which was designed so that learners do not see each other. The recordings start with the first task until the end of the second. Conversations were transcribed and converted to a text document as written data. This was

---

<sup>9</sup> This is based on studies about the tool written in 2012. In 2015, the main intake of learners was still within the age range 26-55, however there is an increase in 18+ students enrolling compared to 2012 (based on UOC internal data).

complemented by researcher notes made during the recordings about learner choices and actions. A screenshot of the task according to task design was also taken. The main researcher, Janine Knight, was familiar with some details of the interface due to former studies where the audioconferencing tool had been used and information about the tool (website and presentation where the tool featured).

The second data set also consisted of recorded oral conversations, approximately 97 minutes of peer-to-peer recorded oral conversations (of 6 cases), of which approximately 53 minutes (before learners changed roles to repeat the task) were transcribed in broad transcription. Recordings were captured using the same plug-in for Skype (first data set). Learners from the second data set uploaded their own recordings in the online classroom. The data starts with the first task until the end of task 2 as both are part of the role-play task. Twenty-two screenshots of the task/interfaces (carried out by the main researcher) were taken and reconstructed chronologically afterwards following the task design. A document explaining the purpose of the task and technological aspects made available to learners online was also collected. Screenshots of 50% of the technological tool logs were also taken (pertaining to the cases chosen). These indicated the task, date, time, number of entries and duration the tool had been open for use. However, screenshots were not taken of all tandem logs because at the point where cases were added to iteratively, some of the log files were no longer available from the online classroom. No log files were available for the first data set because the specific focus on entering/exiting the tool as a navigational act<sup>10</sup> had not yet occurred.

### **8.3 Technological Tool**

The technological tool that learners used for online, synchronous spoken interaction was the Tandem audioconferencing tool <http://www.speakapps.eu/#tandem>. It is a content management application that distributes task materials in real-time and allows students to have different or the

---

<sup>10</sup> The term 'navigational act' is used to refer to the clicking of a navigational screen based resource, that when combined with other clicks, forms a trajectory of use.

same screen-based resources on the interface as they speak. The tool's interface can incorporate photos, text, timers, navigational buttons and pop-ups.

#### **8.4 Tasks**

The tasks were carried out entirely online, without the researcher's presence or intervention. The procedure for the Information gap and opinion-sharing tasks are reported in a study by Appel, Robbins, Moré, and Mullen (2012). In the first data set, two dyads had one interface version and the other two dyads had a different one. As aforementioned, this was for historical reasons because of the origins of the data pertaining to Appel, Robbins, Moré, and Mullen's (2012) study. Both interface versions included the same text-based instructions and photo for partners A and B, as well as navigation buttons or hyperlinks. Each dyad had to collaborate to complete two out of four tasks on the topic of 'travelling'. The first was an Information gap task (spot-the-difference) and the second was an opinion-sharing task. Both tasks incorporated the same photo of a London street. The Tandem tool only distributed materials if both students were connected at the same time, which meant students could not see the materials before connecting with their task partners online. The topic of 'Travelling' corresponded to a unit students had been working on prior to the oral task, and a discussion activity based on the same image used for the first task. Vocabulary, grammar, reading and listening activities had taken place as preparation before the oral task. Learners had four differences to find: the number of streetlights; the number of windows in a building; the colour of a shop awnings and the sky.

In the role-play task the topic was 'The right job' that corresponded to a unit of work of which the spoken interaction tasks formed a part (see 12 below). Prior to the oral task, learners were required to carry out a number of activities pertaining to language development and technological aspects of the activities, which are listed below. In a textual document with the unit instructions that could be found in the online classroom, learners were offered a suggested path for using the materials and carrying out the activities. However, learners could choose to do some steps before another step e.g. install the software (step 9) at step 2, for



example. The procedure for the unit and how the spoken tasks fit within the unit, is now presented. The actions required of learners pertaining to the technology, before starting the spoken interaction task (step 12), is also outlined (steps 9-11).

Learners were required to:

1. write in the forum in relation to the topic, in response to teachers' instructions on the class 'noticeboard'
2. study the topic using online-based interactive materials (focus on reading and responding)
3. find a partner in the forum in preparation to do the spoken interaction task (for step 12)
4. look for an interesting job from the internet and send it to the forum
5. read the other job adverts posted in the forum and apply for the one that they think is right for them
6. study online pronunciation practice materials, watch teachers video and listen to instructions
7. record an oral contribution using Langblog (a tool for oral production)
8. complete a pre-task quiz from the 'resources' section of the online classroom

#### **Prepare for spoken interaction task**

9. download and install skype and install recording software
10. carry out a test call
11. read instructions about how to use the technological tool, Tandem. NB: learners are informed that they cannot go back or do the task more than once in the instructions in the document
12. carry out the pedagogical task divided into two sub-tasks a) ask and answer questions to find out information about a candidate at a job interview b) ask and answer questions about two job offers in order to choose which one is more adequate for them. Read the instructions on the screen as the tasks are carried out.
13. save the recording and post the recording to the forum
14. carry out (optional) speaking tasks using the same Tandem tool

The second task used the same photo, accompanied by an open question: *What kind of activities can people do in a holiday destination like this?* Students received instructions and guidelines and carried out the tasks as compulsory course assignments. There was no time limitation for tasks and task engagement took place in unknown locations and unknown devices<sup>11</sup>.

---

<sup>11</sup> It is reasonable to assume that the spoken interaction tasks occurred in learners' homes, given that the synchronous tasks required a degree of quietness. In addition, evidence that learners were out of the home e.g. traffic noises or explicit references by learners made to external sources were not evident in the audio recordings.

In the second data set, the task was a role-play task (divided into two tasks) in which learners took turns being the interviewer and interviewee. The first task required that the interviewer ask questions to his/her peer in their roles. The second task required the interviewer to describe two jobs that he/she had available on the interface page with the aim that his/her peer decided which one they preferred. After the first two tasks were completed (tasks 1 and 2), the tasks were repeated but the peers changed roles (tasks 3 and 4). The task was timed: limited through a timer that showed the minutes and seconds remaining for learners to complete a task. This was followed by a pop-up that appeared when the pre-determined time was up (4 minutes for task 1 and 7 minutes for task 2). Only the first two tasks were analysed as this was deemed sufficient to answer the research question. The analysis stopped at the point when learners swapped roles and began to repeat the two tasks. This meant that in the data analysis of the task reconstruction, only the perspective of one learner (the interviewer) in the dyads was analysed in relation to the screen-based resources.

## **9. Approach and methods for analysis**

The study took a qualitative approach using a case study method, which is an appropriate method for understanding and exploring a contemporary phenomenon (Yin, 2009). A case study method was chosen because the aim was to understand firstly how learner agency manifested and then as the study evolved, how it intersected with screen-based resources and meaning making. Yin's (2009) 'six steps' were used to guide the research process namely, Plan, Design, Prepare, Collect, Analyse and Share. The approach evolved and changed during the course of the study, which meant returning to the planning and design stage once again. Before presenting the first approach and methods taken, an overview of the data sets, how the data was treated and the papers they relate to, is presented in Table 2. This is then followed by descriptions for each stage of the study. The procedure taken for coding across data sets is presented in section 9.1. and pertains to the original aims of the study to identify if types of agency were evident in other task types. In section 9.2, a description of the new focus taken is presented, and outlines the new approach taken in the

study when it became evident that the screen-based resources were a central part for understanding agency and the shaping of learner talk.

**Table 2. Data sets and data treatment**

<b>FIRST TREATMENT OF DATA SET 1 AS AN EXPLORATORY STUDY</b>		
Data Set 1 4 cases  Information gap and Opinion-sharing	<p><b>Etic perspective</b></p> <p>Coding of meaningful utterances as turns according to three analytical dimensions: 1) cognitive processing; 2) social processing and 3) functional analysis ‘The Analytical Framework of Peer group interaction’ (Kumpulainen and Muntanen, 1999) that incorporated Halliday’s language functions (Halliday and Hassan, 1989) in the functional analysis. Where new functions were identified, they were added to the original coding system. After the data had been attributed to the sub-sections within the analytical dimensions and subsequently coded, another researcher checked the results. Because it had been previously identified that the learners in case 2, 3 and 4 had looked at the answers and scripted or prepared their interaction in some way, we did not fully code within the cognitive processing dimension.</p>	Paper 1  <i>A framework for Learner Agency in online spoken interaction tasks</i>
<b>SECOND TREATMENT OF DATA SET 1 WITH DATA SET 2</b>		
Data Set 1 4 cases  Information gap and Opinion-sharing  and  Data Set 2 4 cases (different from data set 1) Role-play	<p><b>Etic perspective</b></p> <p>Coding across both data sets using ‘The Analytical Framework of Peer group interaction’ (Kumpulainen and Wray, 2002 that is an expanded version on the original framework from 1999, used above). The cognitive and social processing dimensions were not used, only the ‘language functions’ dimensions. Because utterances from the first data set had already been coded, cohesion of coding across the data sets was re-checked. Where new functions were identified, they were added to the language dimensions coding system.</p> <p>Coding of meaningful utterances, as turns, were used using ‘verbal designations’ (Krippendorf, 2013) for defining the semantics of the data. As suggested by Krippendorf (2013), longer definitions were used to describe the language functions used or that emerged in order to avoid abstraction. Examples were put alongside the definitions as a guide for future coding/coders.</p> <p>Data set 2 was chosen on the fact that some cases mentioned the navigational resources and other cases not.</p>	Paper 2  <i>The Negotiation of Shared and Personal meaning making in Spoken Interaction tasks</i>

<b>SECOND TREATMENT OF DATA SET 2 AS AN EXPLORATORY STUDY</b>			
Data Set 2a 4 cases  Role-play	<p><b>Emic perspective</b></p> <p>This data set initially started with four cases after identifying the presence/absence of navigational resources from previously coding (so steps 1-5 had been carried out in the second treatment of data set 1 with data set 2, described previously). The data treatment involved:</p> <ol style="list-style-type: none"> <li>1. Listening to the audio recordings</li> <li>2. Transcription general/note-taking</li> <li>3. Listening/looking for potential cases according to 'good' behaviour and ones that do not follow task design</li> <li>4. Determining potential cases/phenomena (focus on navigational resources and other indicators such as recursivity of speech)</li> <li>5. Looking for recurrent/similar/divergent patterns</li> <li>6. Creating a table of what learners orally make explicit in their audio recordings and adding to the explicit mentions of navigational resources iteratively. The table represented a form of 'tracking'.</li> <li>7. Create more detailed transcriptions</li> <li>8. Identify extracts pertaining to technological aspects including navigation</li> <li>9. Share findings with others</li> </ol>	<p><b>Etic perspective</b></p> <p>Assembling all the data sources (screenshots from researcher task simulation, notes about the classroom (context of case), pdf document with task instructions for learners). Triangulating findings between data sources and checks with course-coordinator.</p> <p>The screen shots were re-assembled chronologically and numbered.</p> <p>The main researcher listened to audio recording while looking at the reconstructed screenshots of the task (according to task design).</p> <p>Tables were made pertaining to similar/different indicators across cases, namely:</p> <ul style="list-style-type: none"> <li>• spoken interaction time in the target language - indicator from data set 1</li> <li>• number of seconds before start of talk-identified as an indicator also in data set 1</li> </ul> <p>Labelling process of the interface pages took place which involved identifying different screen-based resources and labelling the modes associated with them e.g. visual/textual/navigational. The order of the screenshots was then mapped to the task according to design.</p>	<p>Paper 4 and 6</p> <p><i>Navigating a Multimodal ensemble: mediating turns verbally and non-verbally in online interaction</i></p> <p>(paper4)</p> <p>and</p> <p><i>But the computer say me the time is up: the shaping of oral turns mediated with and through the screen</i></p> <p>(paper 6)</p>
Data Set 2b  (2 more cases added from the	<p><b>Emic perspective (focus on textual/visual resources)</b></p>	<p><b>Etic perspective (focus on textual/visual resources)</b></p>	

<p>same sample) Role-play</p>	<p>1. Steps 1 to 5 were repeated, after identifying one case in the original 4 cases that had appeared to make up their own job description. This meant listening to the whole sample again and seeing if this phenomenon emerged again. Two more cases were selected and added to the original 4 cases to make 6 cases in total. 2. Instead of tracking in step 6 above, tables and ticking were made to compare/contrast what each interviewer in the dyad made reference to explicitly to identify what words or phrases from the text they were using or not in their talk</p>	<p>Instances of 'initiations' from the screen were looked for. Text from the screen was analysed to see if learners responded to them and added to the table, e.g. asked the minimum of 5 questions, used textual sample questions, created own questions. In addition, tables were made pertaining to how learners were using the word/phrases on the screen across cases e.g. number of questions asked by interviewer (to look at learners' response to request to act-instruction to ask 5 questions)</p>	
<b>SECOND TREATMENT OF DATA SET 2 WITH DATA SET 1 AS A CROSS CASE ANALYSIS</b>			
<p>Data set 2a and 2b and 1 4 cases and 6 cases Information gap and Opinion-sharing and Role-play</p>	<p>Tracking technique applied to first data set (listening to audio recordings for mention of navigational resources) was applied to first data set. Although this had already been identified in the analysis of data set 1 in the second coding stage described above. This study identified exactly which navigational resources learners' explicit mentions pertained to (alongside other screen-based resources such as the timer).</p>	<p>Data source added to other data sources, namely the learner log files from the technological tool</p>	<p>Paper 5 <i>Multimodal meaning making: navigational acts in an online speaking task</i></p>

## 9.1 Coding process for identifying types of agency

Here the coding, as part of the process for identifying types of agency across data sets, is outlined. The approach taken for coding took the form of 'verbal designations' whereby longer definitions of concepts were used, without abstractions (Krippendorff, 2013). This, according to Krippendorff (2013), allows for freedom in making theoretically motivated distinctions rather than common ones. Interrater reliability was not established due to the change in focus of the study towards learners' mediation with the screen-based resources.

The exploration of how learner agency manifested during spoken interaction used language functions as indicators for different purposes during task processes. A closer analysis of the focus of what the language functions related to during talk revealed a shift in the students' focus during task processes. The results of which are reported in paper 2. This shift in focus spanned across four general areas, namely: a) task management b) task completion (information gap, opinion-sharing or role-play) with their partner c) relating individual self to the task topic and d) co-ordinating the technological aspects of the task. These four foci were where learners made meaning both jointly and/or individually. As a result, different types of purposes were identified during task processes where students employed language as a tool to achieve 'sub-goals' within the task such as to organise themselves to manage the task. These sub-goals are what lead to types and sub-types of agency emerging.

The coding was carried out by the main researcher with the findings shared with a supervisor during and after the process of coding. After carefully examining the data across the three tasks, sixteen individual functions emerged, subsequently labelled as:

1) Interrogative; 2) Responsive; 3) Affective; 4) Informative; 5) Judgemental; 6) Organisational; 7) Reproductional; 8) Experiential; 9) Argumentational; 10) Imaginative; 11) Summative; 12) Confirmative; 13) Opinion; 14) Clarifying; 15)

Greeting or leave taking and a 16) Miscellaneous function. Whereas functions 1 to 10 appeared in Kumpulainen and Wray's (2002) analysis, functions 11 to 16 were added to iteratively.

Examples of four of the functions identified are described below and the entire coding framework developed can be found in the appendix (appendix 2). The examples include at least one example of each function taken from the transcripts. Two points highlighted by Kumpulainen and Wray (2002) also apply to the analysis. Firstly, the identification of each function required attention to much longer stretches of talk than can be given here. In fact, many utterances could not be coded as functions without careful consideration of previous turns and functions. Secondly, a single utterance may be identified as fulfilling more than one function. When this occurred, the function that was deemed the strongest given the utterances before and after were decided on. Thirdly, some functions were present in the form of another such as personal information given in the form of an answer and therefore were classed as such, e.g. A (PI). = Responsive (Answer) (Personal Information).

### ***Interrogative (Question) (Q)***

Questions requiring information were classed as interrogative. There were certain features that tended to distinguish the Interrogative function in the peer-to-peer interaction including words such as 'what', 'how', 'can' or 'have' at the beginning of a sentence like 'Can you see a pole?' and intonation. The use of question words such as 'do' at the beginning of a sentence were not always used to form questions however, possibly because in Spanish or Catalan statements are used instead, as in the second example below. This is a structure in Spanish and Catalan that can transfer to the target language.

*Student A:* How old are you?

*Student A:* I need travel to London?

### ***Responsive (Answer) (A)***

This refers to talk used to respond to a question or statement, e.g.



*Student A:* What colour is the building? Mine is brown.

*Student B:* Yes, brown er between red and brown.

### ***Affective***

The affective function was indicated by the expression of personal feelings and emotions but was noted as being expressed in relation to different foci of the task. The affective function could arise from reassuring a partner, apologising, expressing humour or exhaustion.

*Student*

*A:* wuff.

*Student*

*B:* wuff.

*Student*

*A:* very tough.

*Student A:* I'm sorry, really, really slow it's slowly working.

Related to the topic of the task, it could relate to personal desire or disinterest in the topic in these examples related to geographical places in which the target language is widely spoken.

*Student A:* About the idea of going to go a football camp to watch a football match. I would like to visit the Arsenal stadium, which is my favourite football team.

*Student A:* I prefer working weekend and always I want to move to New York.

### ***Informative (I, PI, IPref, IT)***

When a student used speech as a means of providing information, his/her oral language was classed as serving the Informative function. As with the affective function, the informative function was noted as being expressed at different foci

of the task. Related to the methodological aspects of doing the task, it could arise from giving their partner information that originated from their task instructions and coded as 'I' or from an image that differed from the information their partner had.

*Student A:* I see a lot of people around that corner and I have like two public lights. (describing the image)

Sometimes, the focus was on information about the management of the technology and coded as IT, e.g.

*Student A:* We go to next task?

*Student B:* Ok, huh...but the computer say me that the time is up.

*Student A:* Yes.

*Student B:* So I think, I think Tandem is not working correctly.

*Student A:* No, maybe because I'm in another page, is in the solutions page.

*Student A:* ok. pues ara tenim un problema perque aqui no va gravant.

(Translation: Ok. well now we have a problem because it\* did not record here)

\*the interaction

The function was also used to express personal information (IP), personal preferences, likes and/or dislikes (IPref) in relation to the topic of the task, e.g.

*Student A:* Do you have a drivers' licence?

*Student B:* Yes, I have but I don't like drive.

## **9.2 Methods and approaches in the new focus**

As aforementioned, the new focus arose from the realization that the screen-based resources and different modes appeared to be a central part for understanding agency and how they appeared to be shaping talk. Therefore, research questions and the rationale for the study were revisited and this involved

a number of actions. Firstly, a literature review on audioconferencing, audio graphic conferencing and videoconferencing (task-based and non-task based) for spoken interaction in the SLA literature was carried out in order to identify studies focused on, or that implicitly reported on, non-verbal screen-based resources during spoken interaction or that had explicitly mentioned 'multimodality'. Secondly, reasons in the literature for why there appeared to be a lack of research in this area (mode researched by focusing on screen-based resources rather than mode as a channel of communication) was sought. Other studies on synchronous CMC for language and non-language learning were reviewed, particularly those focusing on oral interaction. A consideration of what methodologies and approaches were available for analysing the phenomenon were also reviewed and in depth reading about multimodality and social semiotics, that fitted with Halliday's (1978) notion of language as being (a) social semiotic (system) and as a means to act and make meaning. This was considered to be the planning stage for carrying out a new exploratory case study. The design stage involved developing a theoretical base with which to explore mode and agency, which resulted in the focus on mediated action. The design followed an embedded, multiple design with replication logic. The second data set was used as the exploratory pilot study and the first sample of data was drawn on, in order to establish if transferability of findings could be established across task types and data sets.

The cases for the exploratory study (data set 2) had already gone through a screening process with cases selected for 'good' behaviour and those that approached the task differently based on indicators from their talk. While it was difficult to ascertain if talk in the role-play task had been scripted or pre-planned (possibly because the interview format is script-like in the turns it requires), the pattern that both data sets had in common was the absence/presence of explicit oral mentions of navigational resources in their talk. This was identified in the coding phase described previously and whereby the code 'IT', as an Informative language function (see appendix), pertained to utterances relating to (negotiating) technological aspects of the task i.e. screen-based navigational resources, rather than headphones falling on the floor for example. The four cases chosen were deemed representative of the sample from the online class.

Multiple sources of evidence were also collected, and a chain of evidence was established. Other researchers reviewed the results and conclusions. These three things supported construct validity. Consent to use data from the online classroom was sought by email and all learners' real names and related information such as where students lived were changed. A case study protocol was created in a secure cloud database, identifying the types of evidence, the documents to be reviewed, review of teacher/students roles in the tasks and task procedures. What was also noted at this point was that the number of seconds before the start of talk appeared to be much longer in some cases than others and this was also true for cases in the information gap task. We hypothesised that this was learners orientating to the screen as they invited or waited for the invite from their partner (reported in paper 4). Also during the listening and re-listening to the audio recordings it was noted that one case had diverged from using the text about the job-descriptions given on the screen. When this finding was listened for in the audio recordings of the four cases already selected and the other dyads from the sample (that had been initially rejected as cases), other dyads were found to have done the same i.e. created their own job descriptions. Because of this, another two cases were added to the four cases originally chosen in the role-play, making a total of 6 cases.

Because the new study used multiple approaches and analytical tools to answer the research questions, each paper is now presented with the methods and approaches outlined for each one.

In Paper 1, a sociocultural discourse analysis approach was used because the main aim was concerned with how the ways shared understanding through talk (as a way of carrying out intentional choices and actions) was developed in social context, over time (the period of the task). Using the Analytical Framework of Peer Group Interaction (Kumpulainen and Mutanen, 1999), oral peer interactions were coded according to three analytical dimensions: the functions of verbal interaction, cognitive processing and social processing.

For paper 2, content analysis was used for studying learner purposes as well as an emic perspective for studying learners' mediation of semiotic resources

through an analysis of their talk. Language functions were used as the coding system from Kumpulainen and Wray's study (2002), (an updated version of Kumpulainen and Mutanen's framework from 1999). This approach was used because the aim was to explore how learners acted with the semiotic resources to negotiate meaning in order to achieve smaller goals across the three task types. With this approach the aim was to know if the codes associated with different types of agency, identified in paper 1, were applicable to different task types or not.

In paper 4 and 6, realizing that there needed to be more of a focus on learners' mediated action with spoken language as a resource (speech as mode) as well as the screen-based resources (especially the somatic mode which appeared to be shaping learners' oral turns quantitatively), notions from the field of social semiotics were incorporated. The approach consisted of taking an emic (learner perspective) and an etic (researcher perspective) on the data. Mediated action was operationalized by drawing on notions from Multimodal (inter)actional analysis developed by Norris (2004) and notions from Conversation analysis developed by Sacks, Schegloff and Jefferson (1974), in particular, the orchestration of turn-taking.

The notions from Multimodal (inter)actional analysis developed by Norris (2004) is an analytical approach that identifies three levels of social (mediated) action: higher-level, lower-level actions and frozen action, each of which deals with a different level of interaction. *Higher-level action* is used to refer to large-scale actions, such as a meeting, and is made up of 'multiplicity of chained lower-level actions' (Norris, 2004). *Lower-level action* is used to refer to smaller-scale actions, for example, gestures or gaze shifts that become chains of lower-level interactions (Norris, 2004). The lower-level actions support the achievement of the higher-level action. Higher-level mediated actions are those actions that social actors usually intend to perform and/or are aware of and/or pay attention to (Norris, 2016). *Frozen actions* are entailed in material objects after the action has taken place, such as the layout of a room. However, because the physical space surrounding the learners' computers could not be observed, this current study did not apply the notion.

Multimodal (inter)actional analysis also deploys the notion of levels of simultaneous awareness/attention namely *foreground*, *mid-ground* and *background*. This is a continuum that facilitates the visual representation of various levels of attention that an individual is simultaneously engaged in (Norris, 2016) whilst completing an action. A person can be engaged in various actions at a particular point in time (e.g. engaging in a research project, Skyping with family members, interacting with girlfriend) (Norris, 2016) and so decreasing attention/awareness can be described by the continuum.

From Conversational analysis (henceforth CA), the principles of 'relevance' and 'orientation' were used. However, the data was not analysed using CA so did not rigorously follow the protocols of CA transcription, since the main focus was on the non-oral sequentiality of initiations and responses including oral turn taking. 'Relevance' was used to identify what screen-based resources were being made explicitly relevant by learners in their talk during the task process. When learners explicitly mentioned a resource orally, this was deemed 'relevant' and was used to track learners across the trajectory of the task process. The notion of 'orientation' was employed in order to focus on learners' oral turns as responses to previous oral turns and learners' potential response/initiations to the screen-based resources. The identification of what a turn's talk was occupied with or the learners' understanding of the previous turn/initiation was, took place. If and how learners "react(ed) to the messages" (Norris, 2006, p.4) was also identified. Alongside this emic perspective, an etic approach was adopted whereby the main researcher worked looking at the screen-shots of the task whilst listening to audio recordings as learners carried out the task. Instances of 'semiotic initiations and responses' (Coffin and Donohue, 2014) (that also encompassed oral turn-taking), from both learner and screen were identified in this way. Learners' physical turns (which can also be initiations or responses with the screen) were operationalised as occurring when learners indicated clicking a screen-based resource. In addition, screen-based turns or initiations were operationalised as when learners responded to them orally, physically or visually (or a combination of these).

Another focus was also on how screen-based resources were shaping speech: quantitatively by measuring the time-on-task of cases, as well as qualitatively by studying the audio transcripts for matches of the screen-based resources in learners' speech e.g. text from the screen.

In paper 5, the aim was to understand how meaning was negotiated non-verbally. For this, Computer mediated discourse analysis (CMDA) (Herring, 2004) was employed to focus on learner talk. In CMDA meaning is studied by focusing on the phenomena of the meaning of words, utterances (speech acts), macrosegments and raising issues as to what the speaker intends and what is accomplished through language (Herring, 2004). Notions from multimodal (inter)actional analysis was also incorporated to understand the relationship between navigational resources, directional agency and tool use.

Findings were shared with supervisors and colleagues for comments and insights, as well as presented at international conferences to gain feedback from peers. Peer-reviewed journal reviewers and editors also reviewed the analysis and conclusions. This process enabled the findings to be validated by other researchers, facilitated alternative perspectives on the data and added to the theoretical framework during the course of the study.

### **9.3 Changes made to the original focus**

After the analysis at the stage of paper 2, it became apparent that the study needed to focus more on learners' mediated action with screen-based resources as well as spoken language, especially the somatic mode. The research questions were reformulated when it became evident that the data required a multimodal approach and that learners were carrying out the goal-directed mediated action of task completion *through* and/or *with* the available multimodal screen-based resources. Therefore, the steps outlined above were by no means linear. Whilst the exploratory pilot study on learner agency, involving the first data set took a sociocultural theoretical perspective to mediated action, it focused solely on speech as the mediational tool. For the 'turn' towards theorizing and analysing mediated action to include learners' use of speech as one of many

modes, a multimodal approach to data collection and analysis was adopted. This meant that the focus of the research changed, as did the methods for analysis, and more data sources were included. Whilst a case study approach for the second data set continued to be the approach, the 'six steps' of the case study process (Yin, 2009) was returned to.

The second data set now became the exploratory study and the first data set was used to identify if replication had occurred across task types and data sets. Much preparation was undertaken to become familiar with, through reading, different theories and approaches to understanding and analysing SLA through a multimodal lens. Ways of understanding and analysing the data multimodally needed to be found. This meant that theoretical perspectives about multimodality and studies pertaining to learners and screen-based resources needed to be built into the original theoretical framework. Even though some studies pointed towards what data might be collected and how it might be analysed, namely from Lamy (2006; 2012), Raudaskoski (1999) and Benson (2015), there was no model that was found that fitted exactly the study's scenario or the focus on agency. This meant that in the preparation stage, what data was collected and how had to be created. For example, eye tracking could not be used because of the challenge of reaching online learners and even if learners did consent, the study was also concerned with the unobservable behaviours of learners during tasks. Despite this, a form of 'tracking'<sup>12</sup> was devised through focusing on the notion of 'relevance' from CA.

Dyads from the second data set were purposively sampled with the same rationale as the first data set. The rationale for sampling was based on selecting those that demonstrated 'good' behaviour according to task design as well as those that approached the task differently. This rationale continued throughout the study but 'good' behaviour specifically evolved to mean a) 'good' behaviour in relation to navigational resources (i.e. following the task design regarding task sequencing) and b) 'good' behaviour in relation to textual/visual resources (i.e.

---

<sup>12</sup> This involved using learner talk as a form of 'tracking', which according to Chun (2013) gives insight into learners' task process.



using the job descriptions given). Cases were added to the existing data set iteratively when saturation point of the findings occurred.

The turn in focus meant that the data sources needed to be expanded to include the materiality of the screen interface. In addition, the unit of analysis was changed from being the dyads and individuals within the dyads to mediated action as the analytical unit between learners and screen.

When the papers were returned from various journals, recommendations were made to include studies from the field of human-computer interaction (paper 6), Conversational analysis and embodied multimodal turn-taking (paper 4) as well as community of practice theory (paper 5). These recommendations by reviewers were integrated into the theoretical framework and/or literature review in the papers.

## **10. Analysis procedure**

A purposive (selective) sampling procedure was carried out in order to select dyads. The dyads were chosen with the criteria that either they followed the task design when carrying out the role-play task or they appeared to have diverged from the original intended design. The data selection took place twice during the study: once when dyads from the role-play task were selected according to indicators that they followed/did not follow task design (based on evidence such as recursivity of speech and speech fragments that would arguably accompany navigational moves). This selection served to answer the research question for paper 2. Dyads were added iteratively and stopped when it was deemed that data saturation had occurred. Incomplete or unintelligible recordings were discarded from the selection, resulting in four dyads being chosen from the role-play task. A second data selection was prompted during the analysis of the four dyads. During this process, the main researcher noticed that some learners were not following task design with respect to using the textual job descriptions on the screen to interview each other. Instead, they had created their own job profiles for the interviews. All audio recordings from the same classroom were listened to repeatedly whilst the main researcher made notes. From this process, two more

dyads were added to the original four cases when data saturation had occurred with this point. This resulted in a total of six cases from the role-play task. It was noted however that with one of the two cases added, further insight was gained in addition to what was learnt from analysing and choosing the original four cases.

The unit of analysis from research question 2 onwards was focused on 'mediated action' (Wertsch, 1998). This is premised on Wertsch's (1998) notion that human action is the most basic unit of analysis in order to describe what people do and why they do it. Thus, the analysis was carried out combining two perspectives of the data: from an emic (learner) and etic (researcher) perspective. The emic perspective focused on what learners' made 'relevant' during oral (inter)action. The etic (researcher) perspective focused on the screen-based resources that learners were engaged with, while taking turns, either as an (oral or non-oral) initiation or response. These two perspectives involved the analysis of different data sources. Analysis of audio recordings, transcriptions and tracking, based on oral interaction between learners in the dyads, supported an emic perspective. The reproduction of the logical trajectory of on screen resources available during the tasks (by sequencing the screenshots) supported an etic perspective. The researcher perspective was also key for gaining familiarity with what was presented on the screen and how the task design 'guided' a learner in task completion. The main researcher also simulated the task process synchronously with a colleague, taking screenshots to document the task process and re-ordering them chronologically. From this, a labelling process of the screen-based resources, (and what appeared on the screen in-between interface pages), was made. These were then categorised following Lamy (2006), who identified screen-based resources as including natural language in its written and spoken forms and visual resources such as icons and images. In our categorisation 'navigational', 'textual' and 'visual' was added, as well as the terms 'static' and 'dynamic' to indicate whether screen-based resources were moving/changing colour or not. The latter two labels were taken from Herring's (2015) rubric of multimodal CMC.

A key term employed in our analysis was 'screen-based resource' (e.g. a visual timer, pop-up on the screen). Where a number of modes made up a resource the

resource was named according to the sequence of modes that learners were intended to mediate with or the 'hierarchy' of the modes important to its designed purpose. For example, a pop-up that appears suddenly needs to be seen first, then read to know what to do with it, and then 'closed' with a physical click. This would be labelled as a visual/textual/navigational resource. Textual information pertaining to the task was labelled as textual/visual, whereas banners designed to signal task sequence were labelled visual/textual. However, this was problematic depending on what mode learners were attending to at any given moment. The interface pages were also labelled according to this logic.

A dual focus was applied for the analysis of mediation: 1) mediated action of the dyad using oral language and; 2) mediated action of the learner/interviewer with the screen-based resources. Only the resources made explicitly relevant, verbally by the learners in the recorded talk were noted. Relevance was evidenced by the use of phrases like "we can confirm" or learners replicating the words on the navigational resource in speech e.g. "next task". Notes were made iteratively, dyad after dyad and were then transformed into a comparative table, used as a cross-case synthesis to compare evidence, identify patterns and see where cases converged and diverged. This contributed to a chronological overview (or tracking) of the trajectories of screen-based resource use (in the part while talking and recording was occurring), as made explicitly relevant by the pairs in their talk. The conversations of the dyads chosen followed general transcription conventions and were segmented according to what was deemed a meaningful utterance and converted to a text document.

A parallel analysis focused on the mediated action of one learner in each case with the screen from an etic perspective. A reconstruction of the learner's 'steps' (the interviewer) was attempted, as if they were following the task design: the main researcher, through listening to the recordings with the screenshots in hand, followed if and how the interviewer responded to the screen-based resources. In this analysis, the interviewers did not necessarily make the screen-based resources verbally relevant as a topic, but they made them relevant by responding to them in some way (whereby the sense making is constructed one-sidedly by students, following Raudaskoski, 1999). This was noted and each case

was tracked for their responses with the results transposed to the table from the first analysis.

Data sources were triangulated with confirmation checks with colleagues about the tasks given and task conditions. Based on these data sources, researcher notes and tables were made in order to compare various similarities and differences between the cases and individual learners' choices relating to how they used the resources.

Now we turn to the procedure for focusing on how talk was shaped with and through mediation (paper 6) with the screen-based resources. The analytical procedure involved the main researcher working to and fro between reading the transcripts and looking at the screenshots to check that there was coherence (or not) between what learners were saying and the natural progression of the designed task as learners should have seen it on the screen. Where words or phrases in the transcripts exactly or partially matched the resources in the screenshots (e.g. the word 'agent' from the screen-based text becomes 'agency' in spoken language) these were highlighted in bold text in the transcripts. Across the dyads, the identification of how each dyad was or was not encountering and using the text in some way was noted. In the analysis, evidence of the resources that were being used, and how learners were using them in their speech, was identified. In this process, the categories were developed iteratively for describing how learners and resources were shaping speech, identifying what learners did/how they used screen-based resources in some way in relation to the content of their turns (the 'what' of the turns).

The notion of 'orientation' from Conversation analysis was used to understand how learners reacted to some of the messages from the screen. In this way, how learners oriented to screen-based resources (the 'what/who' of the turn-taking) was identified. Only the first two tasks were analysed which was sufficient to focus on one learner as initiator/responder in the two tasks presented.

Following this, the analytical procedure for focusing on non-verbal meaning making with the navigational resources (paper 5) is now presented. The

recordings and transcripts of the four dyads (role-play) were revisited to focus more closely on how learners had mediated physically with navigational resources. Therefore, a focus was on learners' explicit oral mentions of the navigational resources that they had made 'relevant', compared to those that actually existed on the screen during the task (as designed). The analysis took a Computer Mediated Discourse Analysis (CMDA) perspective whereby meaning is identified by focusing on the phenomena of the meaning of words, utterances (speech acts), macrosegments and raising issues as to what the speaker intends, what is accomplished through language (Herring, 2004). By using the emic part of the display table from the analysis carried out in paper 4, learners' general navigational trajectory was 'tracked'. In this way, the 'meaning' dimension<sup>13</sup> in CMDA was extended to include how meanings can be made non-verbally and how navigation might shape oral interaction. Notions from multimodal (inter)actional analysis were incorporated to understand the relationship between navigational resources, agency and tool-use.

The analysis for paper 5 also took into account what was learnt from carrying out the process of researcher task simulation as part of the wider study and the analysis of the task interface in the form of the 22 screenshots taken. The researcher simulated the technological part before the pedagogical task of the role-play task (to choose the pedagogical task and to invite a peer technologically) and took screenshots of learners' Tandem tool logs. The scope of this research question brought the navigational requirements preceding the pedagogical task to the fore. This also ensured that the data sources and methods pertaining to the entry and exit to the tool (interpreted as navigational acts<sup>14</sup>) became important.

---

<sup>13</sup> Meaning dimension= Phenomena: meaning of words, utterances (speech acts), macrosegments. Issues: what the speaker intends, what is accomplished through language. Methods: Semantics, Pragmatics.

<sup>14</sup> The intentional clicking of screen-based resource for navigation by humans, that when combined with other clicks, forms a trajectory of use. The concept is being refined (see glossary of terms).

In order to understand the navigational acts as non-verbal meaning-making, the notion of 'Negotiation of Meaning'<sup>15</sup> (as recommended during feedback by journal editors) was incorporated into the theoretical framework of this particular research question. 'Negotiation of Meaning' is a term used by Wenger (1998) as part of the Communities of Social Practice Theory (Lave and Wenger, 1991) that relates to how people experience the world and highlights engagement in social processes as being meaningful. The notion of 'legitimate peripheral participation', also from Lave and Wenger (1991), became useful in the analysis of the data. Results were triangulated following Yin's (2009) recommendations of comparing evidence from different data sources, perspectives and methods. Following this, diverse data sources were used e.g. task instructions (document) and Tandem logs and utilised varying perspectives- learner perspective and researcher perspective plus checks with the course coordinator. In addition, multiple methods were incorporated including listening and note taking, transcribing, researcher task simulation and task reconstruction through sequencing screenshots and listening to audio recordings whilst looking at the screenshots.

## **11. Results and discussion**

Before presenting the results in written form, a summary of the main findings is presented according to each paper in table form (Table. 3).

---

<sup>15</sup> This conceptualization is not the same as 'Negotiation for Meaning' from a cognitivist perspective, although it shares the same term.

**Table 3. Summary of main results**

Papers	Main results
<b>CORE PAPERS</b>	
Paper 1  <i>A Framework for learner agency in online spoken interaction tasks</i>	<ul style="list-style-type: none"> <li>• Learners can use their agency to reconfigure the tasks from spontaneous to planned interaction, with some choices relating to technology impacting detrimentally on interaction time in the target language.</li> <li>• Different tasks were found to filter and channel different types of agency that learners could exercise, namely representational, organizational, and strategic agency as speech acts, and directional agency as a physical act.</li> <li>• Types of agency consisted of different natures and purposes.</li> <li>• The information-gap supported strategic agency and an opinion-sharing task supported personalization and identity construction or representational agency.</li> </ul>
Paper 2  <i>The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks</i>	<ul style="list-style-type: none"> <li>• The foci for meaning negotiation through speech, across tasks, spanned four areas, namely 1) task management; 2) completion of pedagogical task; 3) relating self to task topic and 4) coordinating technological task features.</li> <li>• Learners are multitasking in tasks, attending to different foci at different times or simultaneously.</li> <li>• The online tasks appeared to require a greater need for dyads to negotiate organizational moves as well as the need to attend to technological aspects, including difficulties.</li> <li>• Learners' semiotic budgets appear to be expanded in relation to the technological aspects of the task, which learners acted upon and which shaped how and what was negotiated, offering new opportunities for meaning making.</li> </ul>
Paper 4  <i>Navigating a multi-modal ensemble: mediating turns verbally and non-verbally in online interaction</i>	<ul style="list-style-type: none"> <li>• Learners in a CMC for spoken interaction task mediate with verbal means (spoken language) but also non-verbal means, in the form of screen-based resources in a task.</li> <li>• Audioconferencing, as a channel of communication, is not necessarily "voice, but no image" (Yamada, 2009).</li> <li>• Mediated action is carried out with spoken language as a screen-based resource and textual/visual requests to act (instructions) that require a learner response. The initiation/response turn-taking with these resources supported the purpose of pedagogical task completion as a higher-level action.</li> <li>• Mediated action is also carried out with navigational screen-based resources. The initiation/response turn-taking with these resources supported the purpose of tool management as a higher-level action.</li> <li>• 'Semiotic initiation and response' (Coffin and Donohue, 2014) was a useful notion for characterizing the mediated action for pedagogical task completion and tool management.</li> </ul>

	<ul style="list-style-type: none"> <li>• Lower-level actions come in the form of oral turns (initiation/response) of learners to complete the pedagogical task as a higher-level action. Turn-taking with navigational resources as lower-level actions to support the completed action of tool management was identified.</li> <li>• Multimodal turn-taking was identified as a multisensory process involving oral turn-taking combinations and where learners attended to the screen through their sensory system (visual) and motor system through touch.</li> <li>• Learners initiated/responded to screen-based resources and in turn, some screen-based resources were considered to act and respond to learner action (navigational resources).</li> <li>• Learner agency was manifest through how learners carried out the task through oral, visual and somatic initiations or responses, through the importance and non-importance given to the textual (pedagogical) task instructions and prompts as well as through learners' repurposing of the navigational resources.</li> <li>• The concepts of lower-level and higher-level actions (Norris, 2004; 2006) allowed for analysis of "tools" as semiotic resources for mediation within a technological tool, without losing sight of their interconnection.</li> <li>• Learners' mediational purposes of a tool may differ from a tool's intended design 'affordances'.</li> </ul>
<p>Paper 5</p> <p><i>Multimodal meaning making: navigational acts in an online speaking task</i></p>	<ul style="list-style-type: none"> <li>• Depending on the navigational trajectory of the cases, learners orally negotiated navigational acts (clicking a navigational resource) as part of their peer-to-peer oral negotiation as meta-modal talk or mediated with in such a way that this oral negotiation is eliminated from their talk in the target language.</li> <li>• Different navigational trajectories effected the time spent interacting orally in the target language.</li> <li>• Learners negotiate the purposes of their tool-mediated actions physically through trajectories of use.</li> <li>• The tool's original purpose for learners to follow a pre-designed navigational path is re-purposed and by doing so, the learners' role is no longer that of a tool user but rather, renegotiated to that of a tool manager.</li> <li>• The notion of 'Negotiation of Meaning' from Communities of Social Practice Theory (Lave and Wenger, 1991) offers a way of understanding some learners' navigational choices and acts, to enter and exit the tool, as a way that learners engage in tool use as a social process which is meaningful for them.</li> <li>• The physical re-negotiation of tool purpose can be understood as learners engaging in 'legitimate peripheral participation' (Lave and Wenger, 1991) the process of <i>learning to be</i> and so through the somatic mode involving navigation, learners can negotiate an identity. That is to say make personal meaning regarding who they are commensurate within a community of practice.</li> <li>• Learners' sensory (visual) and motor systems, not just the language system, appear to be what learners use to negotiate in task process. In turn, this negotiation effects task outcomes not only in linguistic terms (spoken</li> </ul>



	interaction time in the target language) but potentially also with respect to forming learners' identity with respect to tool use.
<p>Paper 6</p> <p><i>But the computer say me the time is up: the shaping of oral turns mediated with and through the screen</i></p>	<ul style="list-style-type: none"> <li>• Screen based resources became: 1) embedded or modified in oral turns, 2) resources to initiate and support oral turns; 3) topics of talk whereby learners exchanged oral messages about procedural aspects, task status and the internal state of the learner as identified by Kraut, Fussell and Siegel (2003).</li> <li>• Learners orientated to some screen-based resources as 4) turn takers, such as pop-ups, as if these were participants in the action.</li> <li>• Learners initiated and/or responded to resources orally and somatically revealing how occasionally peer-to-peer talk can resemble a multi-party, multimodal encounter with other human and non-human sources.</li> <li>• Screen-based resources were identified as active or passive agents in physical turn-taking during oral CMC tasks.</li> </ul>
<b>ADDITIONAL PAPERS</b>	
<p>Paper 7</p> <p><i>Navigational acts and discourse: fostering learner agency in computer-assisted language learning</i></p>	<ul style="list-style-type: none"> <li>• Directional agency appears to be present across other CALL scenarios and task types.</li> <li>• Directional agency can be shared across learners and technological features.</li> <li>• There are multiple forms of navigation, including embodied navigation and whole or partial, human body navigation, which can occur in the same task.</li> <li>• Navigational acts can accompany learner-learner talk and also form part of learner-computer 'talk'.</li> <li>• Learners and computers can act as 'semiotic initiators and responders' (Coffin and Donohue, 2014), resembling sequential turn-taking of talk so that both learners and digital technologies can be understood as potential actors in a discourse.</li> </ul>
<p>Paper 3</p> <p><i>Learner agency and its effect on spoken interaction time in the target language</i></p>	<ul style="list-style-type: none"> <li>• Learners used their agency to reconfigure tasks from spontaneous to planned interaction.</li> <li>• Learners exercised their agency in relation to choice and control of technological features through their use of the navigational resources, which afforded them planning time.</li> <li>• Learners exercised their agency in relation to choice and control of language used through code-switching while interacting spontaneously or performing the task completely in the target language, that was covertly planned in some way.</li> <li>• Learners exercised their agency in relation to choice and control of time mode through their use of navigational resources, which afforded them planning time.</li> <li>• The relationship between the three dimensions of agency are mutually supporting.</li> <li>• Learners' ability to reconfigure a) the time mode of the task design; 2) the ways in which technological tools were used and 3) language choice, all impacted on learners' time spent interacting in the target language.</li> </ul>

	<ul style="list-style-type: none"><li>• 'Temporary avoidance' was identified as a potentially emerging communication strategy specific to SCMC spoken interaction tasks.</li></ul>
--	--

A recap of the research and sub-research questions is now presented, accompanied by a summary and discussion of findings from each. This is followed by how the results pertain to other studies and then the final conclusions are outlined.

### **11.1 Results for research question 1**

#### **What are the foci of learners' goal-directed mediated actions in spoken interaction tasks and how do learners negotiate meaning of the content?**

The foci for meaning negotiation through speech across tasks spanned four areas, namely 1) task management; 2) completion of pedagogical task; 3) relating self to task topic and 4) co-ordinating technological task features. Findings revealed how learners' semiotic budgets<sup>16</sup> were expanded in relation to technological aspects of tasks, which learners acted upon and which shaped how and what was negotiated, offering new opportunities for meaning making (paper 2). Specifically, beyond the pedagogical task, the participants in the dyads needed to orally negotiate organisational moves while attending to technological aspects (including difficulties) although whether this was expressed orally depended on the learners' navigational trajectory (paper 2). Learners' oral negotiation of navigational screen-based resources with their peers can be understood as forming part of the meta-modal discourse also found in non task-based, SCMC studies (e.g. Cunningham, Fägersten and Holmsten, 2010; Fägersten, Holmsten and Cunningham, 2010) (paper 5).

With respect to how learners negotiated the content, a number of results were found. Apart from negotiating the content of the task-based interaction through oral turns, learners also negotiated interaction with the screen-based resources through the use of learners' sensory (visual) and motor systems with respect to how and what screen-based resources they wanted to/could use and then how they wanted to/could use them. This was evident across tasks with respect to the

---

<sup>16</sup> 'Semiotic budgets' is a term coined by Van Lier (2000, p.252) that pertains to the language classroom that provides learners with affordances designed to stimulate engagement with the new language. It has since been applied to other language learning scenarios.

textual/visual pedagogical instructions and prompts (role-play), the image of London (opinion-sharing) and the navigational resources (information-gap and role-play).

## **11.2 Results for research question 2**

### **How does meaning making occur with non-verbal modes in online tasks?**

Across the different task types, results suggest that meaning can be constructed and/or co-constructed between 1) learners' sensory system and the textual/visual modes, which becomes evident in their speech (e.g. lexical items on the screen became embedded in their oral turns) and 2) through their motor system and somatic mode (touch) through learners' initiation and response to navigational resources, that involved clicking them. This construction and co-construction of meaning-making can be understood as emerging from mediated interaction (Norris, 2004) with screen-based resources through the process of 'semiotic initiation and response' (Coffin and Donohue, 2014) or multimodal turn-taking.

Regarding the first way in which meaning making occurred with non-verbal modes, the study identified learners encountering and responding to the textual/visual modes of the screen-based pedagogical task instructions and textual prompts using their sensory system (role-play task, paper 4) and screen-based photo (as image) (opinion-sharing task, paper 2). Learners also responded to the screen-based, textual instruction to ask their partner at least 5 questions (role-play) (paper 4). Because all cases asked their partners questions, it was understood that the textual/visual instruction that requested that they ask their partner questions, was a screen-based 'initiation' that needed to be visually and orally responded to. However, the fact that most learners/interviewers asked their partners more than 5 questions (nearer 10 questions), may have been due to the presence of the 11 textual/visual prompts near the instructions. Therefore, although the notion of 'semiotic initiation and response' (Coffin and Donohue, 2014) was used to characterize how meaning was made with non-verbal modes,

the process was not mechanistic or predictable because other resources (in this case visual/textual prompts) also appeared to shape learners' response(s).

Furthermore, individuals in the role-play task (task 2, paper 4) decided which specific textual/visual prompts they would use to construct the questions to ask their peers. In this sense, the oral questions were constructed through the learners' degree of use and non-use of specific lexical items or phrases from the screen-based text. In some cases, the participants rejected the job descriptions given on the screen and created their own jobs. Learners (the interviewers) orally responded to the pedagogical task's textual instruction to describe the two jobs on the screen to their partners. Their description of the jobs to their peer was understood as a semiotic (oral) 'response' to the textual task instruction as a semiotic 'initiation'. Within this however, many learners used the lexical items from the job descriptions (textual) to help them in the construction of their questions. In addition, learners chose the order of what question to ask (selected from the screen-based text or not) and when. With these resources, the meaning-making was constructed one-sidedly by the learner (following Raudaskoski, 1999).

Another example of how meaning making occurred with non-verbal modes is when learners made personal meaning with a photo (opinion-sharing task, paper 1). Many cases identified the photo as being about London without being given this information in the task. Learners shared their desire, lack of desire (to go) and past experiences about London with their partner. Personal and shared meaning was made with their sensory system involving learners looking at and interpreting the image (as mode) in addition to what they brought to it (e.g. their experience and/or desire), which in this scenario was also shared orally with their peer.

The second way in which meaning making occurred with non-verbal modes was through the co-construction that occurred in the relationship between learners' motor/sensory system and the navigational resources. By clicking 'close' on a pop-up (learners respond) and clicking on 'next task' and/or 'see solution' (learners initiate), learners became responders and initiators through the use of

the somatic mode (touch). In addition, the fact that a pop-up dynamically appeared (in order to 'invite' learners visually/textually to 'close' it) indicated that the computer system could also act as a semiotic initiator. Furthermore, the navigational resources responded by moving learners to another interface page when learners initiated a click to 'next task' or 'see solution'. Therefore, meaning making occurred (from the side of the learner), as learners engaged with the navigational resources (with the somatic mode of touch), through their use of their motor and sensory systems because the navigation involved both reading and clicking. Therefore, such navigational acts not only formed part of the human-human discourse by being a topic of talk, but also formed a part of what can be understood as human-computer discourse. That is to say, human-computer initiations/responses, through navigational acts, that resembled the sequential turn-taking of talk. Because of this, both learners and navigational resources were understood as potential turn-takers and therefore potential 'others', as actors in the unfolding meaning-making during the task process (paper 4). Navigational resources were orientated towards as turn-takers in learner talk such as pop-ups that interrupted and presented learners with a message and which learners could respond to orally and/or somatically, revealing how occasionally peer-to-peer talk resembled a multi-party, multimodal 'encounter' (Raudaskoski, 1999). Different navigational resources also acted as 'active agents' (Dourish, Belotti, Mackay and Ma, 1993) or 'passive agents' in physical turn-taking during the task (paper 6). This last result underscored the potential of screen-based resources to become screen-based agents and therefore meaning/sense makers as the computer responded to learners' clicks/initiation by moving learners to another page/place (interface page) through navigation. The results highlight the multi-levelled, non-verbal aspects of meaning-making that can emerge and the important role of learners' sensory and motor systems during task processes (not just the language system).

The act of learners clicking navigational resources that formed part of semiotic initiation and response sequences for meaning making, can be understood as part of multimodal turn-taking more generally between learner-learner and learner-screen in the task-based SCMC scenario (paper 4). Semiotic initiation and/or response was identified as a way to characterise the (inter)actions in the

data and also appeared applicable to other task-based, CALL scenarios involving navigation and talk (paper 7).

The verbal and non-verbal negotiation that occurred can be described through the notion of 'negotiation for action' (Zheng, Young, Wagner, and Brewer, 2009). This notion takes into account the non-verbal meaning making that occurred through mediation with screen-based resources as well as negotiation using spoken language. Both oral and non-verbal 'talk' (human and computer-generated) and the human body were involved in negotiating in many of the tasks (papers 4, 5 and 7). In addition, the notion can accommodate the intentional choices and actions during pedagogical tasks that learners make in the SCMC scenario in relation to the action(s) of task completion. In this current study, these negotiated actions were the pedagogical task completion and tool management.

Whereas, the meaning making process described above can be said to deal with the lower-level (inter)actions (Norris, 2004) involved in pedagogical task completion and navigation for tool management, results also suggested that higher-level meaning making was made non-verbally. Based on the results, we propose that the meaning/s that learners may have drawn from learners physically negotiating how the tool was to be used and not used (higher-level action) through the use of the navigational resources (by deciding and carrying out navigational clicks as lower-level actions) were also being negotiated and therefore made (paper 5). Learners' physical negotiation, through navigation, underscored their engagement in the development of *learning to be* because through the somatic mode of touch they could negotiate an identity. That is to say, make personal meaning as to who they *are*: whether they were tool users or tool managers or indeed another constructed identity in relation to tool use that is commensurate in the online 'community of practice' (Lave and Wenger, 1991). Learners' navigational acts can also be seen as a form of 'contesting' the way the tool has been designed to be used in the online community (paper 5). By creating their own 'lines of desire' (Luckin and Du Boulay, 2003) as navigational pathways that did not align with task design intentions (paper 2), learners may also have been making meaning about how the tool in the 'community' should be used and for what purposes.

### 11.3 Results for research question 3

#### **How does the mediated action with different screen-based resources shape learners' oral turns?**

Mediated action with different screen-based resources were found to shape learners' oral turns differently, both quantitatively and qualitatively. First, how screen-based resources shaped learners' oral turns quantitatively is presented.

Dyads that chose to mediate the task using navigational screen-based resources, in a way that did not follow task design, was found to affect the time spent in the target language negatively (implying reduced oral turn-taking). This was the case for both data sets: learners that carried out an information gap and opinion-sharing task (paper 1) and learners that carried out a role-play task (paper 5).

A textual pedagogical task instruction requested that learners in a role-play task ask at least 5 questions. However, except for one case, many interviewers asked nearer 10 questions, rather than the initial request of 5 (paper 4). The reason for this may have been that the textual instruction to ask at least 5 questions was followed underneath by 11 textual/visual prompts for the interviewer (sample candidate and sample questions). The number of learner oral turns as a response to the instruction may have been shaped more by the number of available textual/visual prompts on the screen than the textual instruction/request to create 5 questions. Therefore, rather than having 'reinforcing roles' (Kress and van Leeuwen, 2001) in the interaction between modes, the 11 prompts may have overridden the textual instruction with respect to the amount of questions that were asked. This is a hypothesis that requires further research.

Now we turn to how screen-based resources were also found to shape learners' oral turns differently in qualitative ways. Results showed that screen-based resources became 1) embedded or modified in oral turns; 2) resources that initiated and supported oral turns and/or 3) topics of talk whereby learners



exchanged oral messages about procedural aspects, task status and the internal state of the learner (Kraut, Fussell, and Siegel (2003) in relation to the resources. Learners were also found to orientate towards some resources as 4) turn-takers, such as pop-ups, as if these resources were participants in the action (paper 6).

#### 11.4 Results for research question 4

##### **How do learners carry out the goal-directed mediated action of task completion *through* and/or *with* the available multimodal screen-based resources? (paper 4)**

With respect to how learners carry out the goal-directed action of task completion *through* and/or *with* the available multimodal screen-based resources, firstly two levels of mediated action were identified, drawing on Norris' (2004; 2006) notions of *lower-level action* and *higher-level action*<sup>17</sup>. A framework applying these two notions to the SCMC scenario emerged, highlighting learners' encounter with different screen-based resources and the actions carried out with them, through 'semiotic initiations and responses' (multimodal turn-taking).

Two *higher-level actions* were identified as large scale actions that learners achieved in the task process. These were the completion of the pedagogical task and the management of the tool. These two *higher-level actions* were made up of a multiplicity of chained *lower-level actions* which were smaller-scale actions. Actions in the mediation process had two distinct but entwined goals and learners' actions appeared to be directed towards these goals.

One *higher-level action*, the pedagogical task completion, was achieved through the *lower-level actions* of verbal turns, jointly co-constructed between peers as well as what can be understood as the response (in the form of a series of turns

---

<sup>17</sup> *Higher-level action* is used to refer to large-scale actions, such as a meeting, and is made up of 'multiplicity of chained lower-level actions' (Norris, 2004). *Lower-level action* is used to refer to smaller-scale actions, for example, gestures or gaze shifts that become chains of lower-level interactions (Norris, 2004). The lower-level actions support the achievement of the higher-level action. Higher-level mediated actions are those actions that social actors usually intend to perform and/or are aware of and/or pay attention to (Norris, 2016).

as the overarching mediated action). The pedagogical task instructions (e.g. create 5 questions) was understood to be the initiation that requested learners to act and learners responded to these by acting (responding to the visual and then acting orally).

A second *higher-level action* of (technological) tool management was achieved through learners' use of the navigational resources as a chain of smaller scale lower-level actions, which were both individually (physically) and jointly accomplished and sometimes orally negotiated. For example, the pop-up initiated a learner response with the message 'Time-up!' which was a visual/textual initiation. Learners had to respond to it, first visually, then physically by clicking to 'close' it. The dynamic visual/textual pop-up can be understood as a screen-based initiation that was responded to by learners somatically. The learners' response of clicking the pop-up led to another response in the form of a change of interface page. This in turn can be understood as a computer response to a learner. The *lower-level action* of navigation meant that learners could reconfigure (navigate back and out of the tool and re-enter, not just navigate forward) and repurpose the original task and tool design (retrieve task information and instructions without being navigated or pressured by potential pop-ups and timers). The framework in figure 1. below illustrates the process described above.

**Figure. 1 Higher and Lower-Level Actions in an Audioconferencing Task**

<b>Higher-level action(s)</b>	<b>PEDAGOGICAL TASK COMPLETION</b>	<b>PEDAGOGICAL TASK COMPLETION</b>	<b>TOOL MANAGEMENT</b>
	<b>Semiotic resource used:</b> Spoken/aural Language <b>Other:</b> Peer	<b>Semiotic resource used:</b> Textual instructions (screen-based) <b>(potential) others(s):</b> a screen-based resource (or combination of screen-based resources)	<b>Semiotic resource(s) used:</b> Navigable resources <b>(potential) other(s):</b> screen-based resource(s) for navigation
<b>Lower-level action(s)</b>	Joint verbal turn-taking	Textual/visual 'request to act' (pedagogical task instructions) that requires a spoken 'response'	Use (and non-use) of navigable resources

To summarise, the textual/visual instructions (middle column) and oral turn-taking (left hand column) supported the completion of the pedagogical task (role-play, information gap and opinion-sharing task) because learners responded to the instructions that requested them to act as well as responding to each others oral turns. The navigational resources (right hand column) facilitated mediation for tool management (learners could move to the next pedagogical task and check their answers, for example) as well as manage the tool in their own way (go back and forth, exiting and entering the tool at different moments and even over different days). The framework highlights that learners were 'multi-tasking' during tasks (paper 2) as learners' carried out oral and somatic turn-taking in response

to various dynamic (moving) and static (stationary) textual/visual screen-based resources. All three mediational means for action often overlapped in the unfolding task-process: they were not as separate as the visual representation of the framework might suggest.

## **12. Conclusion to Main Research Question**

Now the conclusion of the main research question will be presented which to recap was: How does learner agency intersect with screen-based resources in SCMC tasks for spoken interaction and what can we learn about agency and meaning making in this relationship?

Based on the results of the study, the intersection of learners' choices and actions (regarding agency as defined previously), pertaining to screen-based resources in online language learning tasks, can be described as 'systems with tool(s)-mediated goal(s)-directed action(s)'. Meaning-making emerged from this process. This builds on the notion of 'tool-mediated goal-directed action' (Zinchenko, 1985) from a sociocultural perspective.

The systems that mediated the goal of pedagogical task completion (as a higher-level action) involved human systems and the screen-based resources as part of the computer system. The human systems used were the sensory system, motor system (Harré and Gillet, 1994; Bandura, 1999) and/or the language system (Halliday, 1978) that learners used to mediate with some screen-based resources that learners converted to 'tools' for task completion. With these systems, learners carried out their intentional choices and acted on these choices (as lower-level actions) in the higher-level action of task completion.

The system that mediated the goal of tool management (as a higher-level action) also involved human systems (sensory and motor) and the navigational screen-based resources as part of the computer system. The language system (spoken) (human-human) and the text on the navigational buttons (also as part of the language system) was used for navigational purposes. Therefore, verbal and non-verbal turns were co-constructed across humans and across the human-

computer relationship. Turn-taking, as part of the language system, and/or sensory and/or motor system, facilitated mediation of screen-based resources in lower-level actions (such as finding one difference out of five in the task or orally negotiating when to click a navigation button) within the pedagogical task. Examples also include orally responding to an oral comment from their peer and learners responding sensorially/orally to textual/visual pedagogical task instructions to ask questions or somatically, through touch, click on a (visual/textual) pop-up that instructed learners to 'close' it. The lower-level actions that supported the higher-level actions identified above can be understood as being made up of turns formed by semiotic initiations and responses resulting in multimodal turn-taking. However, also with the systems, learners could take turns with screen-based resources if learners or screen-based resources acted as and/or were orientated to as turn-takers i.e. they initiated/responded to a subsequent/previous turn. By learners taking part in a turn in the turn-taking system, the accomplished action of turn-taking was shared across the human and computer system.

The tools were what learners used to accomplish the goal of task completion and included, but were not restricted to, the screen-based resources that learners chose to use for this purpose. Whilst in much of the literature on tasks and technology, 'tools' have been understood as technological tools (such as a program or an application that can run on a computer system or mobile device), the results of this study highlight the importance of what might be considered 'tools within tools'. These are the screen-based resources used as tools that only become relevant as tools for a purpose when learners use them. In this sense, the word 'tools' as plural, indicates the screen-based resources (and the multiple modes that they are made up of) and that may be used for mediated action(s). It is through tool(s) use (screen-based resources and technological tool) that learners negotiated and made meaning.

In this study, mediation was understood as occurring through interaction, specifically, 'distributed oral turn-taking' (Donato, 2000) that can mediate task completion. Based on the results of this study, the notion of turn-taking was extended to include non-verbal turn-taking, initiated or responded to by other non-

verbal modes. Mediation was identified as occurring with the use of spoken language as a tool in the organisation of turn-taking to complete the goal-directed action of pedagogical task completion. Mediation also occurred through learners' interaction with the use of screen-based resources (so becoming tools), which was made evident in the multimodal turn-taking that occurred i.e. the semiotic initiations and responses, whereby the turn sequence of initiation/response indicated the achievement of mediated action. This occurred with different screen-based resources, at different times during tasks and for different purposes so mediation can be understood as mediation(s).

Goal(s)-directed action(s) as plural also underscores the multi-levelled higher-level and lower-level actions that can exist in tasks with technology. The results of the study suggest that there are two overarching goals, not one goal in the task: the higher-level action of pedagogical task completion and the higher-level action of tool management. These were constituted by the lower-level (inter) actions with screen-based resources and oral turns for carrying out smaller goal-directed actions during the task.

With respect to how meaning was made in tasks, the results suggest that meaning emerged from learners' relationship with the tools. That is to say, emerged from their relationship with the 'tools with systems' with respect to what they could do/could not do and chose to do/chose not to do with them. This could be seen in learners' use of language as a tool to carry out different 'smaller' goals during tasks (paper 1 and 2) such as organising themselves. It could also be seen in learners' use of the image in the opinion-sharing task (personal meanings) (paper 1) and the meanings constructed from the use of textual/visual prompts (task 1 role-play) and the job description (role-play task 2). Learners' use of navigational resources to make their own navigational trajectory (paper 2 and 5) is another example. In addition, we propose that meaning making occurred through learners experience with the Tandem tool, such as who they *are* (were) in relation to the 'community of practice' (Lave and Wenger, 1991) (paper 5).

Finally, results suggested that meaning was made in SCMC tasks with modes through which learners either communicated meaning *with* (speech and somatic

mode) or constructed meaning *from* (text, image, speech and also somatic). The use of the screen-based resources that were the material realization of the modes were central to the meaning that emerged in the 'systems with tools' relationship. To conclude, we now turn to the construct of agency, which is the main focus of the study.

Each facet of the 'systems with tool(s)-mediated goal(s)-directed action(s)' can be understood to be inter-related as a whole. When taking a primarily linguistic approach, learner agency was found to be carried out through the human language system using speech. Spoken language (from their peer) was considered a screen-based resource/tool to support the realization of actions e.g. to organize or represent themselves (paper 1). The spoken language (or learner talk) and learner agency was shaped by how learners chose to use the screen-based resources. In turn, screen-based resources shaped learner talk and it was here that types of agency pertaining to talk (paper 1) also emerged.

Using a multimodal approach revealed that agency could be carried out through human systems with screen-based resources (as tools) as part of the computer system. The use of screen-based resources not only shaped learners' talk in different ways, resulting in types of agency emerging with speech (paper 1), but resources could also act as and be orientated to as initiators or responders of learners' turns. Because of this finding, it is conceivable that learners were not the only agents in the scenario (with intentional goals). Screen-based resources appeared to act as agents by semiotically initiating turns or responding to turns and therefore can be understood as part of the sequentiality of turn-taking. Screen-based resources, therefore, may act as or be orientated to (or not), as 'others' (Raudaskoski, 1999). This means that in some SCMC tasks, at least three parties can be considered as being co-present in the task despite the fact that only two people are interacting orally. In this sense, the encounter can be considered to be a potentially multi-partied one.

Agency can be understood as being co-constructed between humans with human systems (oral language) and between humans and computers/digital systems using screen-based resources. This result points towards an understanding of

agency in tasks as systems based and echoes Thorne's (2016) proposal, influenced by Latour (2005):

“that digital environments and the human experience of activity form unified ecologies with agency distributed throughout the system. The possibility of distributed agency does not necessarily imply symmetry between humans and artifacts (see Kaptelinin and Nardi, 2006), but it does suggest that catalysts for action can shift from brains to bodies and to a range of physical and virtual media in the flow of activity. This position contests the dichotomization of artifacts and humans as distinctly independent from one another. Rather, artifacts and humans together create particular morphologies of action.” (Thorne, 2016, p.189).

### **13. How the findings relate to other studies**

With regard to how the use of the screen-based resources shaped task outcome quantitatively, results of the thesis suggests that improved quantity of language production as a benefit of task-based learning (Ziegler, 2016) did not occur for all learners. This was due to different learners' navigational trajectories, that facilitated pre-task planning, which in turn, led to less spoken interaction time in target language. The importance of learner 'process' rather than 'outcome' (Skehan, 1996) appears an important focus in SCMC online tasks, echoing Chun (2013), because spoken interaction cannot be assumed to be recursive and spontaneous.

With regards to how screen-based resources shaped tasks qualitatively, results of the thesis concur with Balaman and Sert (2017a) that the screen-based resources form part of the content of talk and support the direction of on-screen activities, highlighting how the task interface played a major role in the progressivity of talk. While the results of the thesis concur that in order to construct knowledge, learners orientated towards the interface, the results demonstrate that the various specific, screen-based resources on the interface were orientated towards in different ways, and also shaped turns in different ways. Furthermore, results confirm that not only can turn-taking occur trans-modally (Lamy, 2006) between turn-takers/resources but also that the online



nature of the participants' interaction was a relevant part of learners' interactions (Liddicoat, 2010), i.e. the technological resource can 'invite'. This concurs with Raudaskoski's (1999) notion of 'others' in relation to the screen and that some screen-based resources can act as 'active agents' (Dourish, Bellotti, Mackay and Ma, 1993).

The thesis results underscore that audioconferencing is not necessarily "voice, but no image", (Yamada, 2009, p.1). Indeed, screen-based resources can be present and can form part of the content of talk (Balaman and Sert, 2017) which was also identified in non task-based, language learning scenarios by Cunningham, Fägersten and Holmsten, (2010) (pointer function and text box) and Fägersten, Holmsten and Cunningham, (2010) (video, voice chat and whiteboard application), as meta-modal talk. Linked to this finding is the purposes of the content of talk. Whereas in Balaman and Sert's (2017) study, the screen-based resources pertained to the pedagogical task, the results of the thesis highlight that the content of learner talk pertains to both pedagogical task features (Balaman and Sert, 2017) and technological features (Cunningham, Fägersten and Holmsten, 2010; Fägersten, Holmsten and Cunningham, 2010). The presence of talk pertaining to technological task features identified in the thesis, supports the notion that language learners can be faced with an 'expanding semiotic budget' (Blin and Jalkanen, 2014) in digital task-based scenarios. Furthermore, Kraut, Fussell and Siegel's (2003) 'message types' pertaining to doing something physical during collaborative tasks, while talking, were also identified as being applicable to learner talk in this task-based, SCMC scenario.

The thesis results confirmed Ziegler's (2016) assertion, based on a number of studies, that there is a need for task-based researchers to consider the use of technology as an additional task design feature. However, Ziegler (2016) refers to the term 'technology' with respect to 'technological tools' and 'modalities', whereas the results of the thesis show that screen-based resources also shape the online interaction and conceivably could also be considered as task design features. The thesis results concur with Kenning (2010) rather than Ziegler (2016) in that the features of a specific tool, not only the technological tools themselves or the modality as channel (e.g. audio or videoconferencing), should be taken into

account when understanding the complex factors that help shape online interactions. That is to say, the specific screen-based resources could be considered as task features.

While Kurek and Hauck (2014) consider technology to be a means with which tasks are mediated, along with the second language (L2), (i.e. 'mediated twice' Kurek and Hauck, 2014), the thesis results suggest that different screen-based resources should also be considered as mediational 'tools'. The application of Norris' (2004) notions of 'higher-level' and 'lower-level' mediated action (paper 4) gave rise to the notion of 'tools within tools'. This in turn indicated that in addition to tasks being mediated twice, smaller, lower-level mediated actions within tasks can also be considered as being mediated actions.

Regarding learner behaviours pertaining to the tool, the literature provides a whole host of possible factors associated with the behaviours of pre-(pedagogical) task planning and ways of understanding it, including avoidance (Kim, 2014); off screen-behaviour for learning and non-learning purposes (Hampel and Stickler, 2012; Suzuki, 2013) or not matching the learners' time mode preference (Johnson, Aragon and Shaik, 2002), amongst many others (see paper 3). However, other explanations are also viable and to which we now turn to.

The importance of digital literacy and technological familiarity has been raised by a number of researchers including González-Lloret and Ortega (2014), because according to these authors, they may be competencies in need of supporting in technology-mediated tasks. The results of this study, while highlighting the possible reasons for learners planning their oral interactions to different degrees (in paper 2), also concur with González-Lloret and Ortega's (2014) proposal. However, this is only in so far that in addition to the fact that some learners may have needed support regarding the technological tools (potentially for some of the reasons outlined in paper 3), other learners may have intended to simply use the tool in a different way than the task design. This may have been in despite of the design 'affordances' it may have offered. Therefore, while a number of researchers have highlighted the need to consider task design and task-based

learning and teaching in light of the unique affordances and features available in computed-mediated contexts (see Ziegler, 2016 for overview), including SCMC (Ziegler, 2016; Jung, 2016; Hampel, 2006), the thesis results highlight the importance of learners' (multimodal) experience and purposes for the tool. Results suggest that in addition to focusing on how the tool is actually used as a task design consideration (Heift, 2002) learners' purposes are also a design consideration. On the one hand, the decision for some learners not to take up the tool's design 'affordance' to offer spontaneous interaction but instead repurpose the tool with their own attributed 'affordance' might be deemed problematic. However, taking control of the tool for their own purposes and in their own way is also conceivably a desirable manifestation of digital literacy. Furthermore, learners are engaged in many digitally related activities pertaining to various tools before and after the task (e.g. downloading, recording, making a test call, navigating, uploading) alongside other tools within the same unit of work in which the task forms a part (e.g. oral production tools, forums and internet use). Conceivably therefore, learners' 'cultures-of-use'<sup>18</sup> (Thorne, 2016) or how they use other tools 'around' the task may also be a design consideration.

## **14. Limitations and future research**

### **14.1 Limitations**

The study has a number of limitations pertaining to case study research will be subsequently outlined and presented alongside the measures undertaken to overcome these limitations. The specific limitations of the study are also presented.

Limitations for case study research are that the results cannot be generalized to the wider population and that researchers own subjective feelings may influence

---

<sup>18</sup> Historically shaped practices with particular purposes, associations and values that accrue to a tool from its everyday use (Thorne, 2016). Thorne describes various mismatches between designed and preferred tool use amongst students, different purposes and groups of people.

the case study. It is therefore prone to possible bias. Regarding the first limitation, it can be said that rather than generalizability, case study research aims for transferability of findings to other scenarios with similar characteristics. Regarding bias in the identification of types of agency, this was partially overcome by being open to contrary findings (Yin, 2009). In addition, when the preliminary findings emerged (e.g. indicators of types of agency in data collection stage and also in the notion of agency pertaining to the learner) these were also reported to other colleagues to be checked. In addition, engaging in critical dialogue with experts from peer reviewed journals about aspects of some of the results and the theoretical framework built before data collection and after analysis, helped not only to reduce bias but to strengthen external validity.

The external validity (generalizability) of the framework for (inter)action also needs to be validated by future task-based learning researchers as to whether it is transferable or not to other task-based scenarios.

Another limitation of case study research relates to the concern of internal validity although this mainly pertains to the act of making inferences (Yin, 2009). Because the case study is exploratory (and did not seek to explain why learners' chose different navigational paths), in the instances where explanations for learners' behaviours were considered, rival explanations based on a wide literature review were carefully considered (paper 3).

Pattern matching (within and across cases and across data sets) was also used as an analytic tool to strengthen internal validity. Initially this occurred through coding indicators pertaining to agency types from the first data set to the second data set. When the study changed, the pattern of cases that mentioned navigational resources explicitly was identified, strengthening internal validity. Another limitation is that the sample of the first data set was pre-selected before the research questions were formed and specific foci was fully established. This meant that when the sample was used for cross case analysis across data sets, the sample may have been too small. If sample sizes are too small, it runs the risk of the data analysis not reaching saturation point. Despite the fact that it

would have been better to have a larger sample to draw from, the sample that was selected did allow for analytic generalizability to occur.

Regarding methodological limitations on the screen-based resources, whilst both emic/etic perspectives provided fuller insight into both human-human and human-computer interaction, this is not without limitations. One major limitation is that we cannot see whether learners' orientations are happening in real time, verbal reports of their previous actions or whether these orientations co-occur with talk minute-by-minute. However, because the aim of the study is to establish whether mediation with screen-based resources occur in the first place -and if so, which ones, rather than the minute-by-minute account, we take the discourse as a 'fingerprint' alongside the triangulation of sources to be robust enough to answer the research questions. Whilst eye-tracking technologies combined with micro analytic discourse offers an arguably more accurate understanding of the minute-by-minute action, online learners who are not easily reached by such methods or prefer that their pre-recording behaviours remain hidden, may need approaches such as the one used in this study. The incorporation of learner questionnaires about pre-recording behaviours and think aloud protocols, may also increase validity.

## **14.2 Future research**

Regarding future research, five potential areas have been identified: 1) learner agency; 2) language learning with SCMC audioconferencing tools; 3) meaning making as verbal and non-verbal in CMC tasks; 3) turns and turn-taking in SCMC and 4) mode and task-based learning

### **14.2.1 Learner agency**

The role of new types of agency and their relationship with language learning is yet to be determined (paper 1). The construct validity of the different types of agency therefore should be established through inter-rater reliability as a future research area. Then research on whether the types of agency show replicability

across task types could be carried out. Whereas interaction as strategic agency may sustain learners in extended interaction, interaction as representational agency may support other learning benefits such as task personalisation or enhancing motivation (paper 1).

Furthermore, indicators for directional agency appeared to be evident in other current and future CALL scenarios (paper 7). However, the construct needs to be validated, definition more refined more through validation across different tasks that involve navigation and CALL scenarios. If the construct of directional agency is transferable across these scenarios, then further research on it may help to understand how to foster this type of agency in tasks and may also support the goal of digital literacy. Research would also contribute to task design within 'Integrative CALL' that according to Gruba (2004) seeks to foster agency as an aim. In addition, based on the CALL scenarios surveyed in paper 7, future task-based research may need to look more closely at navigation resources as a feature of tasks that learners control (rather than just as a scaffolding tool). This would be in order to understand 'the interplay of modes' (Schnotz, 1999) between the somatic mode and speech as mode so that this feature of tasks might foster optimal language gains in the target language.

Coffin and Donohue's (2014) notion of 'semiotic initiators and responders' (2014) has been key in understanding agency with navigational resources as a shared system between learners and the computer. However, whilst the initiation/response sequence was found to be useful in characterising what we have come to understand as human-human and human-computer multimodal turn-taking as a discourse involving navigation, the sequence needs further research in order to study sequences in greater depth.

Whereas navigational screen-based resources became a particular focus in the study, because of the way their use shaped learner talk, the visual (e.g. photo) and textual resources (e.g. number of prompts for talk) requires further research into what (language) learning benefits they may have. Whereas directional agency, involving tool use, may support digital literacies, fostering representational agency in tasks may have other language learning benefits (or not). The

relationship between a screen-based image and learners' past/future, real/imagined experiences expressed in talk may be a potential means for other learning benefits such as motivation and personalization. Constructs and notions that appear to connect with the unvalidated construct of representational agency include 'imagined communities' (Pavlenko and Norton, 2007) and 'the possible self' applied to language learning (Dörnyei and Ushioda, 2009). In addition, Bakhtin's (1979) notion that people bring their personal biographies, historical and social knowledge and ways of being to all the interactions when they talk, is also relevant. In that sense, 'people are texts' for each other (Erickson and Schultz, 1981).

In addition to types of agency, the 'interplay of modes' (Schnotz, 1999) (speech as mode and different screen-based resources made up of different modes) needs further exploration in order to identify how these resources are orientated to and how they shape talk. Agency understood through the analysis of turn-taking with 'others' (Raudaskoski, 1999) as signs, text/discourse as conversational resources, or entities/conversational agent (paper 4 and 6), has been useful in opening up the possibility of understanding that the screen may be a semiotic field or 'layer' in the action (Goodwin, 2013). However, resources as 'others' may also act as and/or be orientated to by learners in other ways (paper 4 and 6). Given the various orientations that learners may have towards the screen-based resources (paper 6), as well as the intentionality of the resources that may be responded to (e.g. to instruct, to invite, to move learners to another virtual place, paper 4), future task-based researchers might research screen-based resources as being agentic. That is to say, that some resources can act as and/or be orientated to as having an intentionality of their own (paper 7). Following on from this, future research that seeks to analyse agency in CMC tasks might find the conceptualisation of agency as being systems-based useful, where both human and digital systems with technological tools and screen-based resources are used to achieve specific, intentional goals.

#### **14.2.2 Language learning with SCMC audioconferencing tools**

Given that learners can almost completely reconfigure the terms and conditions in which they prefer to and are able to interact because of how they choose to

use the technological tool (paper 1 and 2), more research is needed to identify reasons contributing to the temporary avoidance of spontaneous, synchronous, speaking tasks (paper 3). Contextual, affective, social or individual factors as well as cognitive factors should be taken into consideration. These may be technology or tool-related, and specific to online learners. Based on this information, ways to scaffold learners could then be trialled. The task time mode, however, may not match the learning style of many online learners who have chosen or prefer asynchronicity. If this is the case, this creates a particular challenge for designers of synchronous tasks and which requires further research in order to solve (paper 3).

Finally, future research related to language learning implications have been identified. Results of the study suggest that assessment of task performance as 'process' rather than 'outcome' (the latter proposed by Skehan, 1998, p.95) appears an important focus for online tasks because peer-to-peer interaction cannot be assumed to be in the TL or that naturally occurring talk (recursive and spontaneous) is occurring (paper 1). Therefore, further research is needed to confirm if new interaction types (i.e. passive/active) are occurring in other online tasks.

#### **14.2.3 Meaning making as verbal and non-verbal in SCMC tasks**

Now we turn to the research area of meaning making as verbal and non-verbal in CMC tasks. Results suggest that whilst the construct of Negotiation for Meaning (NfM) in SLA research is important, future research of meaning making in SCMC tasks should take into consideration both verbal and non-verbal ways of making meaning.

Regarding the verbal meaning making, task design may need to take into account the meta-modal discourse that can emerge from interaction with the screen-based resources and prepare learners to use vocabulary they might need to manage them. Future research into whether providing such language scaffolding in order to help learners in dealing successfully with orally negotiating the resources such as making specific words/phrases related to technological



aspects available (e.g. 'timer', 'headphones') is an area for future research particularly if supporting learners maintaining extended recursive interaction in the target language is an aim (paper 2).

Further research is needed into the cognitive challenges posed by an 'expanded semiotic budget' (Blin and Jalkanen, 2014) whereby (pedagogical) input and technological features of tools need to be attended to (paper 2). Cognitive multimodal research and related theories such as 'split-attention effect' that is "the impairment in learning that arises from the need to mentally integrate disparate sources of information" (Sweller, Chandler, Tierney and Cooper (1990) cited in Moreno and Mayer, (1999) p.359) and the notion that learners' attention is a "resource of limited availability" (Wickens, 1984, p.15), may be useful. In addition, the implications of multimedia learning applied to language learning (e.g. Gilakjani, Ismaila and Ahmadi, 2011) is also pertinent. This future research would go towards understanding the impact of learners being both 'semiotic initiators and responders' (Coffin and Donohue, 2014).

Furthermore, navigational acts can be considered a non-verbal way of making meaning as well as having the potential to shape spoken interaction. Navigational acts can also form part of the meta-modal discourse. Therefore, future research that compares different task types through comparing number of turns should take into account that some learners' oral turns may not relate to the pedagogical task but rather, technological aspects. This is important for SLA research focused on Negotiation for Meaning (NfM) (paper 5) particularly in comparisons of which task type facilitates 'more' or 'less' turns.

Given that meaning negotiation is an important research area in SLA theory and research, further research is needed in order to understand how meaning negotiation in CALL scenarios is carried out semiotically involving many modes, not just linguistically. Whilst our study has mainly focused on the somatic mode connected to the human body, future task-based CALL scenarios might also explore the somatic mode in relation to how digital systems can initiate/respond e.g. through "peep" sounds or vibration from QR codes.

#### 14.2.4 Turns and turn-taking in SCMC

Paper 4 highlighted that further research is needed that takes into account the importance of the screen-based resources in spoken interaction tasks (paper 4). Because learners are attending to both their peer and the screen, a purely lingual lens is therefore not sufficient to capture how learners may be positioned as 'semiotic initiators and responders' (Coffin and Donohue, 2014) in the turn-taking process. Conceptualising screen-based resources as 'others' more relevant in their support for turn-taking through non-verbal means (i.e. visually/textually/somatically), by attending to the screen interface's materiality, is therefore important in order to fully understand turns and turn-taking in SCMC<sup>19</sup> to see how and if learners respond to them as turns or not. This is because the analysis of online language learning interaction and orientations to a task interface has largely remained unexplored (Balaman and Sert, 2017). Whilst conversational analysis (CA) appears to be an appropriate approach for identifying learner orientation to the screen (giving insight into the *how* and with *who/what* of the turn-taking sequence), results suggest that combining the CA emic perspective with an etic perspective as different phases in the analysis is not only compatible but facilitates a screen's 'orientation' to the learner. The etic perspective allows for the screen-based resources and modes to also be analysed (if resources act by 'taking' turns and/or if learners respond to them with a turn or not). Future research might attend not only to "the silent visible displays of the hearer work" (Goodwin, 2013, p.8) of participants (in oral interaction) in order to understand the co-operative social organization involved in a single shared action but also to the silent visible displays of the screen (that may or may not also be involved in a shared action). That is to say, multi-party, multimodal accomplishment of the discourse (paper 6).

Future research on turn-taking in SCMC tasks (where screen-based resources form a central part) may find the incorporation of a multi-party participant framework (paper 6) just as other researchers have used a triadic framework (e.g. with computer-based spell checkers- Čekaitė, 2009; Gardner and Levy, 2010;

---

<sup>19</sup> In SCMC task-based scenarios where non-embodied (digital) resources play an important role.

Musk, 2016) in order to account for possible turns made by the screen-based resources and their part, or not, in turn-taking sequences.

Specifically, future studies on spoken tasks where physical activity is occurring simultaneously to learner talk, might look to frameworks from the field of Human Computer Interaction such as Kraut, Fussell and Siegel's framework (2003) when analysing spoken interaction, particularly how the carrying out of simultaneous physical activity can shape talk (paper 6) during a task.

#### **14.2.5 Mode and task-based learning**

Paper 6 touched briefly on the fact that many studies on CMC and multimodality have focused on the use of bi and tri channels of communication (e.g. text chat and videoconferencing) and its impact on spoken interaction. In contrast, this study has focused on mode and semiotic resources, bringing the importance of the screen, rather than the channel to the fore. Within this focus, a number of notions and approaches related to multimodal research, and social semiotics in particular, have been central in understanding and analysing the task-based CMC scenario. In addition to the theoretical notions used to understand the data and form part of the methodological approach ('semiotic initiation and response' - Coffin and Donohue, 2014), multimodal ensemble (Bezemer and Kress, 2016), 'others' as sign/discourse/entities (Raudaskoski, 1999); higher and lower-level (inter)actions (Norris, 2006), other notions have been applicable to the results. Future research on mode in task-based learning might look to already well-established work carried out on multimodality by non-CMC researchers e.g. Kress and van Leeuwen (2001) and later taken up by linguistics including CMC (e.g. Lamy, 2006; Hampel, 2006 and Flewitt, 2008), in order to advance a methodological 'turn' to researching non-embodied screen-based modes and CMC.

The need for conceptual models, based on empirical work, is needed to support design processes and analysis of learning activity in digital spaces for further enhancements (Blin and Jalkanen, 2014) including CMC with screen-based resources. Therefore, the framework for (inter)actions in CMC tasks (paper 4)

which incorporates many notions from social semiotics, needs to be 'validated' as to whether it can be transferable to other similar task-based CMC scenarios.

Future research should focus on the 'interplay of modes' (Schnotz, 1999) including the notion of hierarchy of mode (Kress and van Leeuwen, 2001) in screen-based CMC tasks. This is based on our own results that suggest that in the mediation process, modes can 'override' others (paper 4). The number of textual/visual prompts possibly overrode the textual instructions with respect to the influence on amount of oral turns learners made. This is a tentative conclusion but needs further investigation because the number of visual/textual prompts may be able to support an increase in number of oral turns taken- which, following SLA theory, leads to optimal language learning gains. In addition, the somatic mode pertaining to navigational resources shaped the mode of speech quantitatively and qualitatively (all papers).

From a task-design perspective, the intentionality of different screen-based resources is also highlighted by the results. Whilst some resources serve the purposes of prompting or shaping spoken turns between peers, others are a means for the teacher/designer to communicate information to students whereby no learner response or turn is required, such as technological task instructions rather than pedagogical task instructions. These resources may draw the learner's attention away from their peer, so needs further investigation as to, for example, what and what not to include on the screen. Research on the notion of a foreground– background continuum of attention/awareness (Norris, 2006) may be useful for this.

## **15.Task design considerations**

Regarding navigational screen-based resources from a design perspective, results suggest that while they can be understood as a means to 'scaffold' learners (paper 5), learners may use them as mediational tools for their own purposes. This underscores that "the most important mediator of human behaviour in internet environments is the purpose of the people who visit or inhabit them" (Wallace,1999, p.5). Indeed, learners' purpose/s at the higher-level

(inter)action for tool use seem to emerge in this study as an important design consideration. If learners are using the tool in ways which do not harness the tool's affordance for a specific pedagogical purpose (in this case synchronous, spontaneous spoken interaction for language learning), then in addition to considering adaptations to the lower-level (inter)actions of the navigational resources as scaffolds, task/tool design may focus on the higher-level action of tool management. For example, learner training on how learners can use the tool for different purposes (e.g. planned/scripted spoken interaction vs spontaneous spoken interaction) and the different language learning benefits of each (e.g. less time in the TL vs more time in the TL/ focus on accuracy vs focus on fluency).

Tool design purpose may also need to be rethought in the light of how learners actually use it rather than how it is projected to be used, as suggested by Heift (2002). How learners appear to want to use the tool e.g. as tool managers and/or users is conceivably a task design consideration as well as a focus for future research. For learners to engage in tool use for optimal language gains, issues of digital literacy and 'cultures-of-use' (Thorne, 2016) have implications: conceivably learners not only need to know how to use the technological tools that others design (as digital literacy) but also need to know for what purposes the tools can be used for and the language learning gains that different uses of them support (also conceivably a facet of digital literacy). Tool design for tasks may even consider tool repurposing as a means whereby different language outcomes might be afforded i.e. that planned as well as spontaneous interaction are considered 'legitimate' (Lave and Wenger, 1991) in the online 'community' because this forms part of the everyday culture of how different tools are used for activities carried out around the task.

Connected to tool design considerations are the design considerations for the navigational resources. Whilst navigational resources can be understood with the aim of scaffolding learners as part of task design, as we have seen, learners may use them for their own purpose(s). Learners may even re-purpose the navigational resources from one use to the next use. Therefore, future research regarding navigational resources in tasks might focus on learner purpose and learner tool experience (experience focus) together with the notion of tool

affordance (design focus). This is not only to consider tool possibilities but also learner purposes, which may also highlight a possible misalignment between the two (paper 5). Whether designers should give learners navigational choice and control over his/her own instruction in order to compose a unique and individual instructional sequence (Lawless and Brown, 1997) is perhaps a false dichotomy if learners engage in navigational behaviours that minimise the optimal benefits of a tool so that apparently little language learning gains are made. Focusing on learners' experience of the tool and purposes might help to reframe the issue, as would a focus on who learners *are* (e.g. perceiving themselves as tool users or managers), in relation to the technological tool.

This leads us on to how task design might channel learners' agency so that they use their agency and the tools in ways that offer optimal benefits for language learning i.e. extended synchronous time in the target language. However, as aforementioned learners may not want to use the tool for this purpose. Therefore, the possibilities for what the tool can be used for i.e. planned, semi-planned to synchronous, and the language learning benefits for each planning type may need to enter the discourse between teachers and learners. Whilst on the one hand addressing the scaffolding of tool use (purpose) may involve re-design that considers learners use of it for different types of planning for interaction, task designers may need to concede that ultimately, learners are going to, and if they can, make decisions about tool use for themselves. In addition, and importantly learners' decisions and purposes may be shaped by an array of factors (paper 3), and identifying these is of paramount importance before any changes to task design are made.

Another task design consideration is more concerned with the increasing complexity of how and when tasks are carried out. If network-based learning scenarios are characterised by an expanded range of semiotic resources across multiple sites and time frames, typical spoken interaction tasks (based on the assumption of a synchronous, and static site for interaction) may not fit an online scenario that works predominantly in an asynchronous time mode. Given that current and future tasks are also no longer restricted to a physical classroom (Ziegler, 2016) (or even an online desktop based classroom) and that navigation (e.g. hand or

whole body) might be increasingly a feature of tasks that accompany learner talk (paper 7), how these fit together appears to be an important future task design consideration.

## Bibliography

- Ahearn, L. M. (2001). Language and agency. *Annual review of anthropology*, 30(1), 109-137.
- Appel, Robbins, Moré, & Mullen, (2012). Task and Tool Interface Design for L2 Speaking Interaction Online. In L. Bradley & S. Thouësny (Eds.), *CALL: Using, Learning, Knowing*, EUROCALL Conference, Gothenburg, Sweden, 22-25 August 2012, Proceedings (pp. 15-19). Dublin: Research-publishing.net, 15.doi: [10.14705/rpnet.2012.9781908416032](https://doi.org/10.14705/rpnet.2012.9781908416032).
- Austin, J. L. (1962). *How to Do Things With Words*. Cambridge: Harvard University Press.
- Bakhtin, M. (1979). *Speech Genres and Other Late Essays* trans. Austin: University of Texas Press.
- Balaman, U., & Sert, O. (2017a). The coordination of online L2 interaction and orientations to task interface for epistemic progression. *Journal of Pragmatics*, 115, 115-129.
- Balaman, U., & Sert, O. (2017b). Development of L2 interactional resources for online collaborative task accomplishment. *Computer Assisted Language Learning*, 30(7), 601-630.
- Bandura, A. (1999). A social cognitive theory of personality. In L. Pervin & O. John (Ed.), *Handbook of personality* (2nd ed., pp. 154-196). New York: Guilford Publications.
- Benson, P. (2015). You Tube as text: spoken interaction analysis and digital discourse. In R. H., Jones, A. Chik, & C. A. Hafner, (Eds.), *Discourse and digital practices: doing discourse analysis in the digital age* (pp. 81- 96). Oxford and New York: Routledge.
- Berglund, T. Ö. (2009). Multimodal student interaction online: An ecological perspective. *ReCALL*, 21(2), 186-205.
- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: a social semiotic frame*. London: Routledge.
- Blin, F., & Jalkanen, J. (2014). Designing for language learning: agency and languaging in hybrid environments. *APPLES: journal of applied language studies*, 8(1), 147-170.
- Block, D. (2003). *The social turn in second language acquisition*. Washington, DC: Georgetown University Press.
- Block, D. (2013). Moving beyond Lingualism: Multilingual Embodiment and Multimodality in SLA. In May, S. (Ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 54-77). New York, NY and London, UK: Routledge.
- Breen, M. P. (1987). Learner contributions to task design. In C. Candlin, & D. Murphy (Eds.), *Language learning tasks. Lancaster practical papers in English language education*, Volume 7 (pp. 23–46). Englewood Cliffs, N.J: Prentice Hall International.
- Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 187–206). Cambridge: Cambridge University Press.



- Bueno-Alastuey, M. C. (2013). Interactional feedback in synchronous voice-based computer mediated communication: Effect of dyad. *System*, 41(3), 543-559.
- Calvo-Ferrer, J.R., Melchor-Couto, S. & Jauregui, K. (2016). Multimodality in CALL. *ReCALL*, 28(3), 247-252.
- Candlin, C. (1987). Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language Learning Tasks* (pp. 5-22). Englewood Cliffs N.J: Prentice Hall International.
- Canto, S., de Graaff, R., & Jauregi, K. (2014). Collaborative tasks for negotiation of intercultural meaning in virtual worlds and video-web communication. In M. Gonz alez-Lloret, & L. Ortega (Eds.), *Technology and tasks: Exploring technology-mediated TBLT* (pp. 183-212). Washington, DC: Georgetown University Press.
- Čekaitė, A. (2009). Collaborative corrections with spelling control: Digital resources and peer assistance. *International Journal of Computer Supported Collaborative Learning*, 4, 319-341.
- Chun, D. M. (2013). Contributions of tracking user behavior to SLA research. *CALICO Journal*, 30, 256-262.
- Coffin, C., & Donohue, J. (2014). A language as social semiotic-based approach to teaching and learning in Higher Education, *Language Learning*, 64, Supplement 1, 287-297.
- Collentine, K. (2009). Learner use of holistic language units in multimodal, task-based synchronous computer-mediated communication. *Language Learning & Technology*, 13, 68–87.
- Coughlan, P., & Duff, P. A. (1994). Same task, different activities: Analysis of a SLA task from an activity theory perspective. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 173–193). Norwood, NJ: Ablex
- Cunningham, U. M., Fägersten, K. B., & Holmsten, E. (2010). "Can you hear me, Hanoi?" Compensatory mechanisms employed in synchronous net-based English language learning. *The International Review of Research in Open and Distributed Learning*, 11(1), 161-177.
- Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (Ed.) *Sociocultural theory and second language learning* (pp. 27-50). Oxford: Oxford University Press.
- Dooly, M. (2016). Performing identities in social media: A proposal for studying identity construction in language learning online. In C. Develotte, H. Baldauf, & M. Beldame (Eds.) *Special Issue: Identity construction in social media*, *International Journal of Applied Linguistics*, Vol 20(1). <https://alsic.revues.org/3005> [Retrieved: 22 January 2018].
- Dooly, M. (2015). Networked classrooms and networked minds: Language teaching in a brave new world. In C. J. Jenks & P. Seedhouse (Eds.), *International perspectives on the ELT* (pp. 84-109). London: Palgrave Macmillan.
- Dooly, M., & Hauck, M. (2012). Researching multimodal communicative competence in video and audio telecollaborative encounters. In M. Dooly & R. O'Dowd (Eds.) *Researching online foreign language interaction and exchange: Theories, methods and challenges* (pp. 135-161). Bern: Peter Lang.
- Dornyei, Z. & E. Ushioda (Eds.) (2009). *Motivation, Language Identity and the L2 Self*. Bristol, UK: Multilingual Matters.

- Dourish, P., Bellotti, V., Mackay, W., & Ma, C.-Y. (1993). Information and Context: Lessons from a Study of Two Shared Information Systems. *Proceedings of the ACM Conference on Organisational Computing Systems COOCS'93* (Milpetas, California) (pp 42-51). New York: ACM.
- de la Colina, A.A., & García Mayo, M.P. (2007). Attention to form across collaborative tasks by low-proficiency learners in an efl setting. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 91–116). Clevedon: Multilingual
- Ellis, R. (2000). Task-based research and language pedagogy. *Language teaching research*, 4(3), 193-220.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Erickson, F. & Schultz, J. (1981). When is context? Some issues and methods in the analysis of social competence. In J.L. Green & C. Wallat (Eds.), *Ethnography and Language in Educational Settings*, vol. V (pp. 147-150). Norwood NJ: Ablex.
- European Centre for Modern Languages (ECML), (2017). *Digital literacy for the teaching and learning of languages*. Retrieved from <https://www.ecml.at/Portals/1/5MTP/Ollivier/flyer-elang-EN.pdf?ver=2017-11-22-104100-870>
- Fägersten, K. B., Holmsten, E., & Cunningham, U. (2010). Multimodal communication and meta-modal discourse. In R. Taiwo (Ed.), *The handbook of research on discourse behavior and digital communication: language structures and social interaction* (pp. 144-163). Hershey, PA & New York: Information Science Reference.
- Flewitt, R., (2008). Multimodal literacies. In J. Marsh & E. Hallet (Eds.), *Desirable Literacies: Approaches to Language and Literacy in the Early Years* (pp. 122–139). London: Sage.
- Flewitt, R., Hampel, R., Hauck, M. and Lancaster, L. (2009). What are multimodal data and transcription? In C, Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp.40–53). London: Routledge.
- García, O. (2009). Education, multilingualism and translanguaging in the 21st century. In A. K. Mohanty, M. Panday, R. Phillipson, & T. Skutnabb-Kangas (Eds.), *Multilingual Education for Social justice: Globalizing the local* (pp.140-158). New Delhi: Orient Black Swan.
- Gardner, R., & Levy, M. (2010). The coordination of talk and action in the collaborative construction of a multimodal text. *Journal of Pragmatics*, 42(8), 2189-2203.
- Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA review*, 19(1), 3-17.
- Gilabert, R., Manchón, R., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117-135.
- Gilakjani, A. P., Ismail, H. N., & Ahmadi, S. M. (2011). The Effect of Multimodal Learning Models on Language Teaching and Learning. *Theory & Practice in Language Studies*, 1(10), 1321-1327.
- Gillen, J. (2014). *Digital literacies*. Oxford and New York: Routledge.
- González-Lloret, M. (2011). Conversation analysis of computer-mediated-communication. *Calico Journal*, 28(2), 308–325.

- González-Lloret, M., & Ortega, L. (2014). Towards technology-mediated TBLT: An introduction. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 1–22). Amsterdam, the Netherlands: John Benjamins.
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of pragmatics*, 46(1), 8-23.
- Gruba, G. (2004). Computer assisted language learning (CALL). In A. Davies & C. Elder (Eds.), *Handbook of applied linguistics* (pp. 623-648). Oxford: Blackwell.
- Halliday, M.A.K. (1978). *Language as social semiotic*. London: Edward Arnold.
- Hampel, R. (2006). Rethinking task design for the digital age: A framework for language teaching and learning in a synchronous online environment. *ReCALL*, 18(1), 105–121.
- Hampel, R. (2010). Task design for a virtual learning environment in a distance language course. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 131–153). London, England: Continuum.
- Hampel, R., & Hauck, M. (2006). Computer-mediated language learning: Making meaning in multimodal virtual learning spaces. *The JALT CALL Journal*, 2(2), 3-18.
- Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(02), 116-137.
- Harré, R. & Gillet, G. (1994). *The Discursive Mind*, Sage: London.
- Hauck, M. (2010). The Enactment of Task Design in Telecollaboration 2.0. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology*, (pp. 197-217). London, England: Continuum.
- Heift, T. (2002). Learner control and error correction in ICALL: Browsers, peekers, and adamant. *CALICO Journal*, 19(2), 295–313.
- Hellermann, J., Thorne, S. L., & Fodor, P. (2017). Mobile Reading as Social and Embodied Practice. *Classroom Discourse*, 8(2), 99-121.
- Herring, S. (2015). New Frontiers in Interactive Multimodal Communication. In A. Georgakopoulou & T. Spilioti (Eds.), *The Routledge Handbook of Language and Digital Communication* (pp. 398- 402). New York: Routledge.
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338-376). New York: Cambridge University Press.
- Jenks, C. (2014). *Social interaction in second language chat rooms*. Edinburgh, Scotland: Edinburgh University Press.
- Jepson, K. (2005). Conversations and negotiated interaction in text and voice chatrooms. *Language Learning & Technology*, 9(3), 79-98.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing Multimodality*. London: Routledge.
- Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of interactive learning research*, 11(1), 29-49
- Jung, Y. (2016). *The effects of SCMC modality and task type on negotiation of meaning* (Master Dissertation). Retrieved from [https://englishagenda.britishcouncil.org/sites/default/files/attachments/mda2017\\_special\\_commendation\\_yeonwoo\\_jung\\_university\\_college\\_london.pdf](https://englishagenda.britishcouncil.org/sites/default/files/attachments/mda2017_special_commendation_yeonwoo_jung_university_college_london.pdf)

- Kaptelinin, V., & Nardi, B. (2006). *Acting with technology: Activity theory and interaction design*. Cambridge, MA: MIT Press
- Kenning, M. M. (2010). Differences that make the difference: a study of functionalities in synchronous CMC. *ReCALL*, 22(1), 3-19.
- Kim, H.,Y. (2014). Learning opportunities in synchronous computer-mediated communication and face-to-face interaction, *Computer Assisted Language Learning*, 27(1), 26–43.
- Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-computer interaction*, 18(1), 13-49.
- Kress, G. (2014). What is mode? In C. Jewitt (Ed.), *Routledge Handbook of Multimodal analysis* (pp. 60-75). London: Routledge.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Edward Arnold.
- Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL quarterly*, 34(2), 337-340.
- Krippendorff, K. (2013). *Content Analysis: An Introduction of its Methodology*. Thousand Oaks, CA: Sage Publications.
- Knight, J., & Barbera, E. (2016). The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks. In Pixel (Ed.) *Proceedings of the 9th International Conference on ICT for Language Learning* (pp. 248-252). Italy: Libreriauniversitaria.it
- Knight, J., Barbera, E., & Appel, C. (2017). A framework for learner agency in online spoken interaction tasks. *ReCALL*, 29(3), 276-293.
- Knight, Dooly & Barberà (2018). Navigating a Multimodal Ensemble: mediating turns verbally and non-verbally in online interaction. Manuscript submitted for publication.
- Knight, Dooly & Barberà. (2018). Multimodal meaning making: navigational acts in an online speaking task. Manuscript submitted for publication.
- Knight, Dooly & Barberà. (2018). But the computer say me the time is up: the shaping of oral turns mediated with and through the screen. Manuscript submitted for publication.
- Knight, J., & Barberà, E. (2017). Learner agency and its effect on spoken interaction time in the target language. *JALT CALL Journal*, 13(3).
- Knight, J., & Barberà, E. (in press). Learner agency and its effect on spoken interaction time in the target language. *Electronic Journal of e-Learning*.
- Kohn, K., & Hoffstaedter, P. (2017). Learner agency and non-native speaker identity in pedagogical lingua franca conversations: insights from intercultural telecollaboration in foreign language education. *Computer Assisted Language Learning*, 30(5), 351-367.
- Kumpulainen, K. & Mutanen, M. (1999). The Situated Dynamics of Peer Group Interaction: an introduction to an analytic framework, *Learning and Instruction*, 9(5), 449-473.
- Kumpulainen, K. & Wray, D. (2002). *Classroom Interaction and Social Learning: from theory to practice*. London: Routledge Falmer.
- Kurek, M. & Hauck, M. (2014). Closing the “digital divide” - a framework for multiliteracy training. In L. Williams and J. Pettes Guikema (Eds.), *Digital literacies in foreign*

- language education: Research, perspectives, and best practice*, CALICO Monograph Series, Volume 12 (pp. 119-140). Texas: CALICO.
- Lai, C., & Li, G. (2011). Technology and task-based language teaching: A critical review. *CALICO journal*, 28(2), 498-521.
- Lamy, M-N. (2006). Multimodality in second language conversations online: Looking for a methodology, In A. Baldry & E. Montagna (eds.), *Interdisciplinary perspectives on multimodality: Theory and practice*. Proceedings of Third International Conference on Multimodality, (pp. 237-254). Campobasso: Palladino
- Lamy, M. N. (2012). Click if you want to speak: Reframing CA for research into multimodal conversations in online learning. *International Journal of Virtual and Personal Learning Environments (IJVPLE)*, 3(1), 1-18.
- Latour, B. (2005). *Reassembling the social*. New York: Oxford University Press.
- Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional science*, 25(2), 117-131.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, UK.
- Lave, J., & Wenger, E. (2002). Legitimate peripheral participation in communities of practice. In M. Lea, & K. Nicoll (Eds.), *Distributed learning: Social and cultural approaches to practice* (pp. 56-63). London: Routledge.
- Liddicoat, A. J. (2010). Enacting participation: Hybrid modalities in on-line video conversation. In C. Develotte, R. Kern, & M. N. Lamy (eds.), *Décrire la conversation en ligne*. (pp. 37–50). Lyon, France: ENS Éditions.
- Liou, H. C. (2011). Blogging, Collaborative Writing, and Multimodal Literacy in an EFL Context. In M. Levy, F. Blin, C. B. Siskin, & O. Takeuchi, (Eds.), *WorldCALL International Perspectives on Computer Assisted Language Learning*. (pp. 3-18). New York: Routledge.
- Luckin, R., & Du Boulay, B. (2003). Lines of desire: the challenges of interactive educational TV. *Proceedings of the 11th International conference on artificial intelligence in education (AIED 2003)* <http://rp-www.cs.usyd.edu.au/~aied/papers.html>
- Martin, J. (2004). Self regulated learning, social cognitive theory and agency. *Educational Psychologist*, 39(2): 135-145.
- Mercer, S. (2012). The complexity of learner agency. *Apples – Journal of Applied Language Studies*, 6(2), 41–59.
- Miller, E. (2012, Sep 2). Chapter proposals: Agency and SLA. Retrieved from <https://linguistlist.org/issues/23/23-3666.html>
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of educational psychology*, 91(2), 358.
- Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse studies*, 9(2), 194-225.
- Mondada, L. (2013). Embodied and spatial resources for turn-taking in institutional multi-party interactions: Participatory democracy debates. *Journal of Pragmatics*, 46(1), 39-68.
- Musk, N. (2016). Correcting spellings in second language learners' computer-assisted collaborative writing. *Classroom Discourse*, 7(1), 36-57.
- Musk, N. (in press) 'Chained' spelling corrections in English learners' collaborative writing. *Discourse Processes*.



- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. London: Routledge.
- Norris, S. (2006). Multiparty interaction: a multimodal perspective on relevance. *Discourse studies*, 8(3), 401-421.
- Norris, S. (2016). Concepts in multimodal discourse analysis with examples from video conferencing, *Yearbook of the Poznan Linguistic Meeting*, 2(1), 141-165. doi: <https://doi.org/10.1515/yplm-2016-0007>
- Novick, D. G & Sutton, S. (1997). What is mixed-initiative interaction. In the working notes of the AAIL spring symposium on computational models for mixed initiative interaction (pp. 114-116). AAIL/MIT Press.
- Pavlenko, A. & J. Lantolf. (2000). Second language learning as participation and the (re) construction of selves. In J. Lantolf (Ed.) *Sociocultural Theory and Second Language Learning* (pp. 155–77). New York: Oxford University Press.
- Pavlenko, A., & Norton, B. (2007). Imagined communities, identity, and English language learning. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 669-680). New York: Springer.
- Raudaskoski, P. (1999). *The use of communicative resources in language technology environments. A conversation analytic approach to semiosis at computer media* (Doctoral dissertation). Retrieved from <http://vbn.aau.dk/ws/files/72280035/pirkkosphd.pdf>
- Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning, supplement 61*(s1), 1-36.
- Roebuck, R. (2000). Subjects speak out: How learners position themselves in a psycholinguistic task. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 79–96). New York: Oxford University Press.
- Rossolatos, G. (2015). Taking the “multimodal turn” in interpreting consumption experiences. *Consumption Markets & Culture*, 18(5), 427-446.
- Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation, *Language*, 50(4):696-735.
- Satar, H. M., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, 92(4), 595-613.
- Sauro, S. (2009). Strategic use of modality during synchronous CMC. *CALICO Journal*, 27(1), 101-117.
- Sauro, S. (2011). SCMC for SLA: A research synthesis. *Calico Journal*, 28(2), 369-391.
- Schnotz, W. (1999). Introduction [Special issue]. *European Journal of Psychology of Education*(14), 163-165.
- Schwartz, D., & Okita, S. (2009). *The productive agency in learning to learning by teaching*. Retrieved from May 30, 2018 from [http://aaalab.stanford.edu/papersProductive\\_Agency\\_in\\_Learning\\_by\\_Teaching.pdf](http://aaalab.stanford.edu/papersProductive_Agency_in_Learning_by_Teaching.pdf)
- Searle, J. N. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, England: Cambridge University Press.
- Sinkovics, R. & E. Alfoldi (2012). Facilitating the interaction between theory and data in qualitative research using CAQDAS. In G. Symon & C. Cassell (eds.), *Organizational Qualitative Research: Core Methods and Current Challenges*. (pp.109-131). London: Sage.

- Skehan, P. (1992). Strategies in second language acquisition. In *Thames Valley University Working Papers in English Language Teaching*, no. 1.
- Skehan, P. (1994). A framework for the implementation of task-based learning. In *Thames Valley University Working Papers in English Language Teaching*, no. 3.
- Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 17-30). Oxford, UK: Heinemann.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Stake, Robert E. (1981). The art of progressive focusing, *65th Annual Meeting of the American Educational Research Association*. Los Angeles.
- Stein, P. (2004). Representation, rights, and resources: Multimodal pedagogies in the language and literacy classroom. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 95-115). Cambridge, England: Cambridge University Press.
- Stickler, U., & Shi, L. (2017). Eyetracking methodology in SCMC: A tool for empowering learning and teaching. *ReCALL*, 29(2), 160-177.
- Stockwell, G. (2010). Effects of multimodality in computer-mediated communication tasks. In M. Thomas, & H. Reinders (Eds.), *Task-based language learning and teaching with technology*. London: Continuum International Publishers Group.
- Suzuki, S. (2013). Private turns: A student's off-screen behaviors during synchronous online Japanese instruction. *CALICO Journal*, 30(3), 371-392.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J.P. Lantolf (ed.) *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.
- Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced Language Learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176-192.
- Sykes, J. (2005). Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO*, 22, 399–432.
- Thomas, M. and Reinders, H. (Eds.) (2010). *Task-Based Language Learning and Teaching with Technology*. London: Continuum.
- Thorne, S. L. (2008). Computer-mediated communication. In N. Van Duesen-Scholl & N. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 4. Second and foreign language education* (2nd ed., pp. 325-336). New York: Springer.
- Thorne, S. L. (2016). Cultures-of-use and morphologies of communicative action. *Language Learning & Technology*, 20(2), 185-191.
- van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. Lantolf (Ed.), *Sociocultural theory and second language acquisition* (pp. 245-259). Oxford, England: Oxford University Press.

- van Lier, L. (2008). Agency in the classroom. In J. P. Lantolf and M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 163– 186). London: Equinox.
- Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied linguistics*, 6(1), 71-90.
- Vetter, A., & Chanier, T. (2006). Supporting oral production for professional purposes in synchronous communication with heterogenous learners. *ReCALL*, 18(1), 5-23.
- Vitanova, G. (2005). *Authoring the self in a non-native language: A dialogic approach to agency and subjectivity*. In J. K. Hall, G. Vitanova, & L. Marchenkova (Eds.), *Dialogue with Bakhtin on second and foreign language learning: New perspectives* (pp. 149–169). Mahwah, NJ: Lawrence Erlbaum.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wallace, P. (1999). *The psychology of the Internet*. New York: Cambridge University Press.
- Wang, J. (1996). *Same task: different activities*. Unpublished research report. Pittsburgh, PA: University of Pittsburgh.
- Wang, Y. (2006). Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL*, 18(1), 122–145.
- Warschauer, M. & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching*. 31, 57–71.
- Warschauer, M. (2000). *CALL for the 21st Century*, Presented at the IATEFL and ESADE Conference. July 2, 2000, Barcelona, Spain.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: Cambridge University Press.
- Wertsch J. V. (1998). *Mind as Action*. Oxford University Press, Oxford.
- Wertsch, J.V. (1991). *Voices of the mind*. Cambridge, MA: Harvard University Press.
- Wickens, C.D. (1984). *Engineering psychology and human performance*. Columbus, OH: Merrill Publishing.
- Yamada, M. (2009). The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: Experimental study. *Computers & Education*, 52(4), 820-833.
- Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: It's about time! *Language Learning & Technology*, 14(3), 72-93.
- Yin, R. K. (2009). *Case study research: Design and methods (4th ed.)*. Thousand Oaks, CA: SAGE.
- Zheng, D., Young, M. F., Wagner, M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal*, 93(4), 489-511.
- Ziegler, N. (2016). Taking technology to task: Technology-mediated TBLT, performance, and production. *Annual Review of Applied Linguistics*, 36, 136-163.
- Zinchenko, V.P. (1985). Vygotsky's ideas about units for the analysis of mind. In J.V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 94—118). New York: Cambridge University Press.





## **17. Articles in journals or conference proceedings**

- 17.1 Paper 1: A Framework for Learner Agency for Spoken interaction tasks

# *A framework for learner agency in online spoken interaction tasks*

JANINE KNIGHT

*Universitat Internacional de Catalunya, Spain*  
(email: [janine@uic.es](mailto:janine@uic.es))

*Universitat Oberta de Catalunya, Spain*  
(email: [jknight@uoc.edu](mailto:jknight@uoc.edu))

ELENA BARBERA

*Universitat Oberta de Catalunya, Spain*  
(email: [ebarbera@uoc.edu](mailto:ebarbera@uoc.edu))

CHRISTINE APPEL

*Universitat Oberta de Catalunya, Spain*  
(email: [mappel@uoc.edu](mailto:mappel@uoc.edu))

---

## Abstract

Learner agency, the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives (Martin, 2004), is instrumental in second language learning because attainment is only arrived at by learner choice (Pavlenko & Lantolf, 2000). If attainment is understood as learner engagement in synchronous, collaborative, spoken interaction which is thought to lead to gains in second language acquisition (SLA), then design considerations that harness learners' agency towards that end is important. This study explores the relationship between learner agency and two different task types, namely an information-gap task and an opinion-sharing task in two peer-to-peer synchronous computer-mediated communication (SCMC) spoken interaction events. Students' choices and how students act on these choices during tasks are analysed using a discourse analysis approach. Audio recordings of four dyads as cases were examined using three analytical dimensions: language functions of verbal interaction, cognitive processing and social processing. The results show that most learners used their agency to reconfigure the tasks from spontaneous to planned interaction, with some choices and actions relating to technology impacting detrimentally on interaction time in the target language. The different tasks were found to filter and channel different types of agency that learners could exercise, namely representational, organisational, and strategic agency as speech acts, and directional agency as a physical act. These types consisted of different natures and purposes and are presented as a framework. The information-gap task supported strategic agency and an opinion-sharing task supported personalisation and identity construction or representational agency.

Keywords: learner agency, task-based learning and teaching, TBLT, synchronous computer-mediated communication, SCMC, spoken interaction, task design

---

## 1 Introduction

The ubiquity of technology in everyday life as well as the many digital environments inhabited for work, play or socialisation, increases opportunities for languaging about the world and language as we engage in diverse activities (Blin & Jalkanen, 2014). New digital contexts are increasingly able to facilitate the construction of personal and shared meanings because the multimodal inputs or semiotic resources and the ways they are configured have expanded, offering new opportunities for learner choice and control or agency. This means new challenges for online task design for language learning. In online synchronous computer-mediated communication (SCMC) milieus, task-based learning and teaching (TBLT), one of the most common design approaches in second language acquisition (SLA) has migrated from face-to-face to online contexts (see Lai & Li, 2011, for an overview). TBLT approaches can be used by designers to stimulate peer-to-peer spoken interaction with the goal of increasing time spent in the target language (TL) and provide opportunities for learner engagement in using a range of language functions.

On the one hand, tasks might be considered to facilitate learner agency, giving opportunities for learners to use the TL in extended interaction and therefore promoting SLA (Samuda & Bygate, 2008). Learners are assumed to have control over their linguistic and cognitive resources, choosing which ones to use for task completion. On the other hand, because some perspectives consider a task as an essentially predictive device, in some cases deterministic (Ellis, 2000), the controlling nature or deterministic features of some tasks might constrain agency.

Whereas some researchers call for tasks to be “less structured, more enquiry-based spaces [that] encourage(s) learners to exercise agency and enact identities” (Lamy, 2006: 263), others warn against the use of highly open-ended tasks (e.g. Doughty & Long, 2003; Lafford & Lafford, 2005), recommending instead to follow TBLT design principles in which meaning is primary, there is a real-world relationship, task completion has priority, and task performance is assessed in terms of outcome (Skehan, 1996). Some propose a shift in design because new learning takes place across networks, multiple sites and timescales, and students have potential agency to create new situated activities (Blin & Jalkanen, 2014). How typical tasks such as information gaps fit within this digital learning landscape is unclear as they assume a fixed, constant site for learning and a synchronous time mode (e.g. a face-to-face classroom), without considering how the presence of technological task features and the ability for learners to control them might shape the interaction.

Learning activities can be organised to enhance agency for two reasons: (a) agency has inherent value both within and beyond formal learning environments; and (b) it leads to superior learning (Schwartz & Okita, 2009). The value is in the belief that learners are agents who “play a defining role in shaping the qualities of their learning” (Dewaele, 2009: 638), and have goals, motives and intentions (Dörnyei & Ushioda, 2009). Superior learning can be considered to be that which fits with a learner’s life goals because it is personally meaningful. Studying how agency manifests in online contexts is important for a number of reasons: firstly, to identify how learners’ intentional choices and actions support the goal of learner “attainment”, understood as the act of interacting orally with others in a spontaneous and sustained way in an additional language; secondly, the need to verify that the peer-to-peer spoken interaction occurring is of the kind that is believed to lead to gains in acquisition; and thirdly, to explore how new forms of agency are emerging in digital

landscapes in order to inform future task design. This study explores how agency manifests during two different pedagogical spoken interaction tasks, online. The results form a tentative framework for types of agency.

## 2 Theoretical framework

### 2.1 *Learner agency, language learning and speech*

There is a plethora of definitions regarding learner agency. Martin's (2004: 135) definition as "the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives" is preferred over the more common "the socio-culturally mediated capacity to act" (Ahearn, 2001: 112). This is because, although many agree that an individual's capacity to act is socioculturally, contextually and interpersonally mediated (Mercer, 2011), Martin's definition allows for a focus on agency as intentional behaviour during task processes. It goes beyond "capacity" or "capability", which are conceptualisations of agency as a property of the individual.

Agency is a fundamental construct in understanding learning processes and learner identities, but researching it is problematic: the construct is under-theorised, there is a lack of clarity in defining it, difficulty exists in establishing sound analytic research procedures, and operationalising it remains a challenge (Miller, 2012). Many perspectives agree that language plays a central role, with a shared focus on learners' individually constructed and renegotiated relationships with society at large (e.g. various authors in Deters, Gao, Vitanova, & Miller, 2015) including online courses or classrooms (e.g. Xiao, 2014; Zhang, 2010; Yim, 2011). At the level of speech, Blin and Jalkanen (2014) connect agency and "linguaging" (Swain, 2006), highlighting a shift in the concept of "language-as-object of study" to "language-as-action or process" involving "making meaning and shaping knowledge and experience through language" (Swain 2006: 98).

Few studies have focused on agency at the micro level of tasks involving speech. In face-to-face contexts, van Lier (2008) identified four areas of learner initiative: (1) topic work; (2) selection to speak; (3) allocation of speaker(s) or activity; and (4) sequencing the talk or activity. Novick and Sutton (1997), focusing on choices, identified choice of outcome – i.e. "a purposeful pursuit of a particular goal", "choice of task" and "choice of speaker". Both studies highlight that agency can be carried out at two levels in tasks: in relation to the topic (i.e. topic work); and in task management (e.g. selection, allocation, and sequencing). Others have focused on how learners construct the same task through different activity: approaching tasks differently despite having the same instructions (Wang, 1996) or with respect to maintaining the frame of social interaction (a set of shared expectations as to what the interaction ought to entail) and positioning themselves differently (e.g. subjects in an experiment or students in a university – Roebuck, 2000); or having different perceptions of and orientations to task conditions across-learner and within-learner performance (Coughlan & Duff, 1994).

A framework for analysing agency at the level of speech does not currently exist in SLA theory. Based on the results of this study, we take Schwartz and Okita's (2009) three types of agency as a starting point. Their study on learning-by-teaching identifies types of agency as: (a) doing; (b) productive; and (c) passive/active. The last two are pertinent because they focus on interaction between people, whereas the first applies to individuals.

Productive agency is “a recursive system [...] where people can create themselves in the world and see themselves reflected back through the independent behaviour of their creation” (Schwartz & Okita, 2009: 8). An example given is a mother and child interacting where the child initiates the exchange, including an idea, into the joint space. The mother incorporates the child’s intent and takes the initiative to turn the conversation into an object-naming lesson. The child picks up the mother’s meaning and renames the object. If the mother never heard the child’s utterance, then the child would have no agency in that interaction, despite choosing to speak. However, the mother takes the child’s idea and builds upon it, which creates a system of productive agency: (a) there is an opportunity to express one’s original ideas; (b) other people take up the ideas and add their own element to them; and (c) one gets to respond with new ideas in return (Schwartz & Okita, 2009). In contrast, passive/active agency can be understood as learners giving ideas while passively receiving others’ ideas during interaction with little or no take-up of what the other is saying. An example given by Schwartz and Okita (2009) is interacting with people who add nothing to the conversation or simply say “yes” or “no.” Interaction is without exchange, only serving to trigger self-explanation. It is semi-recursive.

The characteristics of the spoken interaction in this study differ in two ways to the examples above. The first is that it is peer-to-peer student interaction, not parent-to-child. Peer-to-peer interaction differs because it offers an equalising of participation structures: the authority source (teacher or parent) is subverted as they become a participant in the interaction; control of and responsibility for the interaction is more incumbent on students and speakers to share the floor more equally (Ortega, 1997). The second difference is that the interaction is between non-native rather than native speakers, which, according to Ortega (1997), provides a non-threatening forum for practising developing language skills. However, despite differences in participant characteristics and the features of the speech, compared to Schwartz and Okita’s (2009) example, recursive interaction is desirable for both native and non-native speakers. It is an indicator of “highly engaging conversations” (p. 11) of native speakers and is another way of describing extended and collaborative turn-taking, important for SLA. Productive agency (or recursive speech) and passive/active agency (or semi-recursive speech) remain relevant in studying non-native peer-to-peer interaction in online tasks because, as the data reveals, it highlights that not all spoken interaction involves spontaneous, natural, turn-taking. These two types of agency are the starting points for describing how spoken interaction differs in the online tasks analysed. By developing the two types further we hope to describe the interaction occurring and inform future task design and research.

## 2.2 *Tasks, task types and spoken interaction*

Breen (1987) highlighted the difference between “task-as-workplan” (concerned with the expectations and intentions of task design) and “task-as-process” (or what actually happens), underscoring “the notion that learners, as active agents in learning processes, can modify activities according to their own intentions – modifications which may or may not be in direct accordance with the initial intentions of that task-as-workplan” (Dooley, 2011: 72). Important to this study, some learners’ choices and subsequent actions to reconfigure aspects of the tasks were understood as students’ agency at work, highlighting resistance to do the task-as-workplan (e.g. which tools, when and how).

Much research from a cognitivist perspective has tried to determine which task types have a positive effect on quantity of “meaning negotiation” through turn-taking (Foster & Ohta, 2005), with research suggesting that information-gap tasks offer most opportunities for negotiation and therefore time in the TL. This is important because online task design for spoken interaction often aims to encourage quantity of interaction between learners alongside the assumed recursivity that spontaneous interaction implies. Interaction (involving collaborative, recursive, turn-taking) is central in SLA theory because it is believed to lead to gains in the TL (de la Colina & Mayo, 2007; Gass & Mackey, 2006).

### **3 General objectives**

This study aims to explore how agency manifests within TB-SCMC spoken interaction tasks: specifically (a) how learners use language-as-action (linguaging) for intentional purposes, alongside other means; (b) how jointly-constructed speech as a form of agency is constructed alongside individual agency (the latter being most prominent in sociocultural perspectives – Swain & Deters, 2007); (c) how agency manifests at “task topic” and “task management” levels, extending van Lier’s (2008) and Novick and Sutton’s (1997) research in contrast to classroom or course level or “society at large”; and (d) how task types might induce, “filter” (Luckin, 2010), enable, or constrain agency. With these foci, the study differs from previous ones from a sociocultural perspective. The research questions are as follows.

How does learner agency manifest in:

RQ1: a synchronous online spoken interaction event?

RQ2: an information-gap task and an opinion-sharing task at task management level?

RQ3: an information-gap task and an opinion-sharing task at topic level?

## **4 Method**

### **4.1 Participants**

The participants were students enrolled in English as a foreign language classes as part of their degree programme at the Universitat Oberta de Catalunya (UOC), a 100% online university based in Barcelona. The eight participants comprised three male and five female, non-native speaking adults, aged between 26 and 55 years old, and who were engaged in an online synchronous speaking task as four dyads. All learners were considered bilingual, sharing Catalan and Spanish. Students had a global level B2.1 (upper-intermediate) in English on the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001), although the specific spoken interaction level for each participant is unknown.

### **4.2 Context**

The data derived from a previous study by Appel, Robbins, Moré and Mullen (2012). They took a cognitivist perspective on spoken interaction and used a quantitative analysis of turn-taking and student questionnaires, in order to explore how different navigation features of the Tandem tool (Appel, Nic Giolla Mhichíl, Jager & Prizel-Kania, 2014) influenced learning strategies and TL use in tasks. The results showed that learners engaged in different

activities depending on the various technologies used to do the tasks. One format allowing for easy access to content prior to the task led most learners to prepare or script their interaction, with a considerable amount of preparation for the task but little spontaneous conversation during it. On the other hand, when students had had no prior access to the contents, the task yielded natural spontaneous speech with a higher rate of turn-taking. The current study uses a qualitative approach from a sociocultural perspective in order to gain insight into learning processes.

Four dyads participated in this study: two from one interface version and two from another. Both versions included the same text-based instructions and photo for partners A and B, as well as navigation buttons or hyperlinks. Each dyad had to collaborate to complete two out of four tasks on the topic of travelling. The first was an information-gap task (spot-the-difference) and the second was an opinion-sharing task. Both tasks incorporated the same photo of a London street. Learners had four differences to find: the number of streetlights; the number of windows in a building; the colour of a shop's awnings; and the colour of the sky. The second task used the same photo, accompanied by an open question: "What kind of activities can people do in a holiday destination like this?" Students received instructions and guidelines and carried out the tasks as compulsory course assignments. There was no time limitation for tasks and task engagement took place in unknown locations, although typically it was completed in learners' homes.

The data consists of recorded spoken conversations between four dyads, approximately 23 minutes in total. Recordings were captured using a plug-in for Skype, a free video and audio conferencing tool, and started with the first task and lasted until the end of the second. Conversations were transcribed and converted to a text document, along with researcher notes about learner choices and actions, to form written data.

In terms of instruments, the Analytical Framework of Peer Group Interaction by Kumpulainen and Mutanen (1999) was used to analyse the interactions. It supported a microanalysis of evolving peer interactions, focusing on three analytical dimensions: the functions of verbal interaction, cognitive processing, and social processing. The functional analysis supported investigation of what learners are "doing" with language, and incorporates Halliday's language functions (Halliday & Hasan, 1989) using codes such as Informative, Expository, and Organisational. The analyses of cognitive and social processing focused on interactive dynamics as they occurred across the participants: cognitive processing focused on ways in which students approached and processed learning tasks, highlighting working strategies and situated positions towards knowledge, learning and themselves as problem solvers (e.g. exploratory speech or procedural speech); social processing characterised the social relationship and types of participation in peer groups (e.g. collaborative, individualistic).

### 4.3 Procedure

The study used a *sociocultural* discourse analysis approach concerned with how agency through "content, function, and the ways shared understanding is developed, in social context, over time" (Mercer, 2010: 9). This is a departure from *linguistic* discourse analysis approaches more commonly used in SLA, which are concerned with the organisational structure of spoken language. It also adopted a qualitative approach for data collection and analysis. Time spent in interaction in the TL is included as an important result. Where new



language functions were identified, they were added to the original coding system. After the data had been attributed to the sub-sections within the analytic dimensions and subsequently coded, two other researchers checked the results. Because we know that the learners in cases 2, 3 and 4 had looked at the answers and scripted or prepared their interaction, we did not fully code within the cognitive processing dimension. The speech units were identified on an utterance basis (Kumpulainen & Mutanen, 1999). The unit of analysis was each dyad and the individuals within that dyad.

## 5 Results and discussion

This section describes the results, followed by the presentation of a framework for types of agency identified in the data and how they pertain to productive and passive/active types of agency.

### RQ 1. How does learner agency manifest in a synchronous online spoken interaction event?

Learner agency manifested in two ways: (a) physically through touch, intersecting with technological task features; and (b) through learner speech and spoken interaction, intersecting with methodological features. Learners' relationship with the methodological and technological features of the two tasks influenced the type of agency they could exercise and therefore task features became a specific focus during analysis.

#### Learner agency as physical interaction with technological task features

We consider learners' choices and physical moves made in relation to technological task features of the Tandem tool to be the first type of agency. Although physical in nature, it also implies a spoken or written aspect accompanying or preceding it, necessary for decision-making. The choice and physical moves to navigate (with a button or hyperlink), check or submit answers individually or collectively, using non-device-related technological features, we call "directional agency". Directional agency emerges from the learners' relationship with technological features and affects both task outcomes and processes, specifically the time interacting in the TL and whether the type of spoken interaction is recursive or not, as shown in Table 1.

Case 1 followed task-as-workplan, looking at answers as a checking mechanism after finding their own. Cases 2, 3 and 4 looked at the answers before engaging in interaction in the TL and scripted or pre-prepared what to say before recording, appearing to complete the task.

Table 1 *The effect of directional agency on learners' spoken interaction*

Use of agency to view answers	Spoken interaction time in the TL			Interaction type
	Total	Information gap	Opinion sharing	
Case 1 Answers not seen	11'41"+0'2" in L1	9'15"	2'28"	Recursive
Case 2 Answers seen	5'7"	2'18"	2'49"	Semi-recursive
Case 3 Answers seen	2'27"	1'38"	0'49"	Semi-recursive
Case 4 Answers seen	4'33"	3'34"	0'59"	Semi-recursive

Cases 2, 3 and 4 also used an uncommon word (*awnings*), confirming they had found the meaning before recording. The different choices and moves made by learners, as part of learners' directional agency, led to recursive or semi-recursive interaction. This "split" occurred because case 1 carried out both cognitive (problem solving) and social processing (collaboration) spontaneously during the task. Cases 2, 3 and 4, negated the cognitive dimension by looking at the answers, so social processing consisted not of interacting in the TL to solve a problem but rather interacting/performing the task in the TL as if they had solved it. Case 1 was successful because their choice and subsequent moves resulted in an increased quantity of spoken interaction in the TL, yet despite extensive collaboration they were unable to find all the differences. The other cases, however, appear successful at task completion.

Both quantity and quality of interaction is affected if learners use their directional agency to look at answers before starting the task. Spoken interaction time in the TL is shorter and lacks the recursive quality at both cognitive and social levels. This result highlights the importance of task processes over task completion, as learners who completed the task "correctly" did not engage in spontaneous, recursive interaction.

### **Agency manifest through learner speech and spoken interaction**

After analysing learners' relationship with the technological features, we now describe the relationship between learners and methodological task features. Alongside the closed question (a feature of the information-gap task) and the open question (a feature of the opinion-sharing task), the photo was identified as affecting the types of agency learners could exercise. Learners' choices and their ability to act through speech were filtered, channelled and influenced by methodological task features which led to the emergence of "organisational agency", "strategic agency" and "representational agency", which we subsequently illustrate. These types of agency are speech-related in nature and, due to learners looking at answers and preparing, resulted in a split into either "productive" or "passive/active" sub-types.

### **RQ2. How does learner agency manifest in an information-gap task and an opinion-sharing task at task management level?**

#### **Information-gap task**

Agency in the information-gap task manifested at the level of speech, characterised by the use of language functions for organisational purposes, coded as "OR". When combined together, learner turns formed what we call organisational "moves", or organisational agency. These moves were either: (a) collaborative (COLL); (b) individualistically executed, some of which may have been pre-decided collaboratively before recording (IND/COLL); or (c) individually expressed as self-talk (ST) or intrapersonal speech (e.g. reviewing task requirements). Although all cases demonstrated organisational agency, only case 1 used collaborative moves as shown in Example 2, with one learner in case 1 also using self-talk to organise as shown in Example 1. Cases 2, 3 and 4 mainly used individualistic moves.

#### *Example 1: Individual move as self-talk (ST) (case 1)*

H. OK... (mumbles reading the exact instructions given in the task)... you can see the same picture. There are four differences...

*Example 2: Collaborative move (COLL) (case 1)*

M. ¿Pues empezamos en inglés? (trans. So shall we start in English?)

H. ok. I'm ready to start in English

M. ok. Me too

H. um er. Who a start? You or me?

M. You can start if you want

H. ok

Case 1 demonstrated a greater range of organisational moves, number of turns to make a move, and more time spent in negotiating for organisational purposes than other dyads. Case 1 also codeswitches between Spanish and the TL as shown in Example 2, negotiating when to start speaking in the TL, a move absent in other cases. The use of the learners' shared L1 supported case 1's ability to organise and start the task in the L2, enabling organisational agency to occur across different language codes. This highlighted code change as a natural expression of learner agency (García, 2009).

In cases 2, 3 and 4, organisational agency is manifest at both social and individual levels but it has little purpose as most organisational decisions have been pre-decided. Individual learners in cases 2 and 3, for example, initiate talk but continue without negotiation (agreement/disagreement) of the move with their partner. Cases 3 and 4 announce their roles individually without negotiating them in the TL. These results suggest that learners' looking at answers, as part of their directional agency, detrimentally affected their organisational agency with regard to the range and quantity of organisational moves they could negotiate. This resulted in a passive/active sub-type of organisational agency emerging, characterised by short, semi-recursive interaction with little or no negotiation of organisational moves in the TL.

Table 2 summarises each organisational move and whether it was collaboratively negotiated in each dyad (COLL), individually expressed as self-talk (ST) or individualistically executed (IND). "No" is used when the move is absent.

**Opinion-sharing task**

The opinion-sharing task followed the information-gap task so the purpose of organisational moves was different (i.e., supporting task transition). However, learners' organisational agency was still evident in the form of collaborative, individual (self-talk) and

Table 2 *Organisational moves in the information gap task*

Learner 'move'	Case 1	Case 2	Case 3	Case 4
Choosing when to change to TL	COLL	no	no	no
Choosing when to start	COLL	IND	IND	no
Clarification of what next task involves	ST	no	no	IND
Re-capping what needs to be done for task completion	ST	no	no	no
Choosing sequence of speakers	COLL	no	no	no
Choosing picture, role A or B	–	–	IND*	IND*

\* choosing roles – feature of html version only

Table 3 *Organisational moves in the opinion-sharing task*

Learner move	Case 1	Case 2	Case 3	Case 4
Prompt for next task	COLL	no	IND	no
Clarification of what next task involves	ST	no	no	IND
Choosing when to start	no	IND	no	no
Choosing sequence of speakers	no	COLL	no	no
Choosing when to finish	COLL	COLL	no (recording cut)	COLL

individually executed moves. Cases 2, 3 and 4 continued to have a smaller range of organisational moves than case 1 and were less collaborative in the TL. Examples include learner initiation of an organisational move but executing it on behalf of the dyad or instructing their partner to start, as shown in Examples 3 and 4, resulting in passive/active organisational agency. A summary of moves is presented in Table 3.

*Example 3: Individualistic move (IND) (case 4)*

- A. and then we have to talk about activities that people do when they are in holidays in London and when I am there I always go to shopping because there are a lot of flea markets and I love markets such as Camden Market and Notting Hill as well
- B. so I like to go to museums to theatres and take a lot of pictures

*Example 4: Individualistic move (IND) (case 2)*

- D. Now if you want we can start with task number two so you can start
- J. OK. About task 2 I want to say that I would like to travel to London...

*Example 5: Individual move as self-talk (ST) (case 1)*

- M. yes ok what is next task? (very low voice) Ok. next task (raised voice) Yes, so now we have to discuss
- C. about

In both tasks learners could select to speak and sequence the talk (van Lier, 2008) but could not allocate the speaker roles (A or B) as these were distributed to them via the Tandem tool.

### **RQ3. How does learner agency manifest in an information-gap task and an opinion-sharing task at topic level?**

#### **Information-gap task**

Two factors intersected with learner agency at topic level. The first was the methodological task features that channelled learner agency into interaction for strategic purposes. The second was how the physical move of looking at answers impacted on spoken interaction.

At topic level, case 1 used language function combinations as a strategy to solve a problem, which we call strategic agency. Cases 2, 3 and 4 used combinations to appear to

solve it, using passive/active strategic agency. The information-gap task ensured quantity of interaction (case 1) rather than supporting meaning making derived from the topic. The closed question affected the type of agency exercised because learners could not discuss the topic of “travelling” or implied subtopics from the photo (e.g. “London” or “a busy city street”) so did not reflect choices related to “topic work” (van Lier, 2008). Instead, their agency was channelled towards strategy use to solve the problem, revealing the deterministic nature of the task. This task, with a closed question and hidden information, supported strategic agency. Strategic agency was jointly constructed by dyads using simple to complex combinations of language functions in attempts to find the differences which is shown in Table 4. Authentic attempts were made to find each difference (case 1) or appearing to attempt to find them (cases 2, 3 and 4). Table 4 illustrates this process and also indicates if cases were successful in spotting a difference (hit) or unsuccessful (miss) and the total number of attempts. Misses and hits were determined when dyads started describing a new object in the photo.

Out of the four cases, case 1 made many attempts, including many unsuccessful ones. They also demonstrated the widest range of strategy use (language function combinations) resulting in more interaction time in the TL. While case 1 demonstrated exploratory interaction at cognitive level (making attempts by using complex language function combinations), cases 2, 3 and 4 engaged in cognitive processing, appearing to explore. However, because the problem-solving element of the original task had been removed, cases 2, 3 and 4 engaged in passive/active strategy use characterised by: minimal hits/lack of attempts beyond the number of differences (four); reduced complexity of language function combinations; a limited range of language function use (predominantly information-based swapping strategy use, e.g. I+I); and fewer instances of using language functions for collaborative problem solving (e.g., question/answers and/or agreement/disagreement). The dyad’s strategic agency resulted in productive or passive/active agency sub-types and these were determined by learners looking (or not) at the answers, using their directional agency in the first place.

### **Opinion-sharing task**

The opinion-sharing task resulted in diverse learner interpretations and responses to the topic both across and within the dyads. The photo of a London street combined with an open question channelled learners’ engagement into personal meaning making. Learners expressed feelings and/or related their experiences “to personalise and otherwise enrich what is to be learned”, an example of agentic engagement (Reeve & Tseng, 2011: 258). Cases 2, 3 and 4 connected their experiences with the photo (focusing on London as the topic), situated themselves in the past (recounting past experiences of London) or the future (desire to go to London), expressed their feelings about London, or reasoned why it was good to go there. This highlighted how individual students approached the same task differently (Wang, 1996; Roebuck, 2000) and how the task type permitted cases 2, 3 and 4 to be “imagined agents” (Anderson, 1983) or real agents who are living/re-living the experience as speakers of English, with personal meanings and social purpose in a place in which their TL predominates. Learners could construct “multiple identities involved in the process of learning and using an L2” (Swain & Deters, 2007: 821) through the perceived topic. This possibility is “filtered” (Luckin, 2010) out in the information-gap task. In all instances, learners personalised the topic through the selection, control and representation of personal meaning with respect to

Table 4 Language function combination for strategy use

Language function combinations	Breakdown of attempts: Hits (H) and Misses (M)	Total attempts
Case 1 (I+AF) + (I+RE) + (Ja/Jd) + I + Ja I + (I+Jd) + I + I + I + (I+Jd) + I + Q + A (I+Q) + strategy interrupted because of a technical problem + I + I + Ja + I (Q+I) + (Ja+I) + Ja + I + Ja + I + Ja + I + Ja + I (Ja+I) + (Ja+I) + Ja + I + Ja I + I + Ja + I + (Ja+Jd+I) + (Jd+I) + I + I + Jd + RE + Ja I + I + I + Q + A + Ja + Q + Cl + I + Ja + I + Q + Cl + (RP+Q) + Q + Ja + A + (Ja+I) + RE + AF I + (RP+I) + I + I + (Ja+I) (I+Q) + Ja + I + (RP+Jd) + RE + RE + RE + AF + (RE/OR) Q + I + Ja + I + RP + I + Ja + I + Q + I	M H (number of street lights) M (car sign) H (building colour) M (metro signs) H (building/windows) H (number of windows) M (rows of windows) H (awnings) M did not find last difference (sky)	10
Case 2 I + (I+RE) + (RE+I) + RE + Ja I + (RP+Q+Q) + A + (I+RE+Sum) I + (I+RE) + Ja (RE+I) + I + (RE+Sum)	H (sky) H (number of windows) H (awnings) H (street lights)	4
Case 3 I + (Ja+Sum) I + (Ja+Sum) (I+Sum) (I+Q) + I	H (street lights) H (number of windows) H (sky) H (awnings)	4
Case 4 I + I + RE + Ja I + I + (RE+Sum) (I+Q) + (A+I) + I + I + Q + A I + I + (I+Sum) (I+Q) + (A+I) + Sum Q + A + I/Q + Jd + (RE+Sum) + Ja	not exploratory – describing H (street lights) not exploratory – describing H (awnings) H (sky) H (number of windows)	4

Key to language function codes: providing information (I); expression of personal feelings (AF); reasoning in oral language (RE); answering questions (Q); reproducing spoken language (RP); answering questions (A); organising behaviour (OR); summarising (Sum). Functions in brackets indicate the combination of language functions used in 1 learner utterance; + without brackets indicates the next speaker.

Codes from Kumpulainen and Muntanen (1999) with the exception of “Sum”.

their own experiences, likes, interests, lack of interest or desires. We call this representational agency.

Despite learners' personalisation of the task, cases 2, 3 and 4's interaction was mainly individualistic, not taken up by their peers, resulting in semi-recursive interaction (see Example 6). Passivity by partners was evident until their turn (silent, no interruptions, no questions, uptake of ideas or acknowledgement of ideas without listening). Contributions were characterised by a lack of cohesion in ideas and different foci within the topic,

resulting in the interaction as a whole not making sense. On occasions, speakers did not share the floor more equally or share being the authority figure – usual features associated with peer-to-peer interaction (Ortega, 1997). We call this type of interaction passive/active representational agency. This individualistic delivery could be explained by one or more factors: learner avoidance, communication anxiety, individuals preparing the answers differently (with agreed topic but without agreed shared orientation), or having different things to share as individuals.

*Example 6: case 2*

- D. now if you want we can start with task number two so you can start
- J. ok er about the task two um I want to say that I would like travel London er it's a special city with beautiful place and monuments there mixture of human race and is one of the most important cities of the world. I would like to go London for learn English. I think that this is good reason for visit London. Also in London you can do many things for example you can learn English, meet new people or friends, walk for several parks or see alone the Thames River in a boat. Also you can show clothes and souvenirs, visit the London Tower and the British Museum and if you like sports er you can see a football match and after in the evening you can drink a delicious beer in a typical English pub. What do you think?
- D. yes I totally agree with you Joel. I think that London is a very cosmopolitan place.
- J. uhuh
- D. It has a lot of places to visit like Big Ben... in Buckingham palace
- J. uhuh
- D. and I also guess it must be a nice place for going out for a drink with many nightclubs it would be nice to have a pint of Tetleys
- J. uhuh
- D. and London also have many places to go shopping like Harrods.
- J. uhuh ok
- D. About the idea of going to go a football camp to watch a football match. I would like to visit the Arsenal stadium, which is my favorite football team.
- J. ok
- D. um Also it would be nice to visit Wimbledon tennis courts
- J. ok
- D. it must be a nice place where Rafa Nadal is winning every year.
- J. ok. I agree

In contrast, case 1 (see Example 7), do not agree a topic of talk as they move between discussing “a place” to “shopping”. Partner M shares geographical and cultural information unique to her through focusing on shopping but partner H appears uninterested in both the place and his partner’s talk of shopping. Partners appear to interpret the open question differently: one as a problem like the information-gap task that preceded it, and the other as a discussion point, highlighting differences in task perception and orientation (Coughlan & Duff, 1994). Their recursiveness is short-lived; possible explanations include lack of motivation, shared interest or knowledge of the topic, tiredness or different interpretations of task requirements.

*Example 7: case 1*

H. maybe make or doing Tai chi

M. what

H. take the metro the underground to go to some place some beautiful place to make Tai Chi I don't know what can you do in a place like that

M. After you know that yesterday was the black Friday it's a very big shopping day here

H. ah yeah

M. so I can only think about going shopping

H. yes it's the only thing you can do. Yes I don't know

M. I don't know I guess a place like that that you can go with your friends or your family and then

H. ah ok and then you can go to the pub

Recursive interaction in tasks with open questions might only be able to be maintained by natural drivers. Interaction time in the TL is less across all dyads in this task compared to the information gap.

Table 5 summarises the types of agency identified in the data and sub-types as productive and passive/active, taken from Schwartz and Okita (2009).

## 6 Conclusions

Results suggest that learners can exercise their agency through speech, using language as an action (linguaging) or through physical moves. These two natures of agency were combined during task processes to carry out learner intentions. Whereas the physical moves were related to technological features, learner acts using speech were related to the learners' relationship with the methodological task features and the need for task management. The learners collaborated using speech to organise themselves (organisational), solve a problem (strategic) or represent themselves (representational), all of which had intentionality. This suggests that agency is not a single, monolithic factor (Mercer, 2011) but rather non-monolithic, multi-natured and complex (e.g. types of agency can occur simultaneously or sequentially).

In addition, the results highlight the importance of how both technological and methodological features of tasks can shape learners' spoken and physical acts during interaction but can also be shaped by learners in the mediation process. Learners' relationship with features can filter, channel and influence the types of agency that they can enact. A physical learner move involving technology was detrimental to time spent in the TL, indicating that if interaction time in the TL is the objective, complete learner control in relation to technology is not always desirable. The importance of a dual-focused application of agency in research, which is both causal-focused and rights-focused, is therefore important. The former relating to "protecting or enabling people's access to a particular form or expression of learning" and the latter "concerns what conditions foster learning, and not what rights are at play" (Schwartz & Okita, 2009: 7).

## 7 Implications of the study

Digital environments may expand the types of agency that can be exercised. How learners exercise these types may or may not support spontaneous, recursive, spoken interaction,



Table 5 *Types and sub-types of agency*

Type and nature of agency	Sub-type	Definition	Characteristics
<b>Directional agency</b> Physical interaction with technology accompanied by spoken interaction	none	Choices preceding and accompanying physical moves made by learners to e.g. navigate, check or submit answers either individually or collectively using non-device related technological features	<b>Individual (physical) interaction</b> Joint negotiation (speech/written)
<b>Organisational agency</b> Spoken	Productive	Negotiating organisational moves prior, during and/or at the end of a task	<b>Recursive and collaborative</b> A wide range of negotiated moves relating to task organisation Evidence of intrapersonal speech
	Passive/active	Performing organisational moves that have been pre-decided in the L1 and not negotiated in the TL	<b>Semi or non-recursive</b> Individualistic delivery Limited range of or absence of necessary organisational moves
<b>Strategic agency</b> Spoken	Productive	The use of language function combinations to solve a problem collaboratively	<b>Recursive and collaborative</b> Wide range of language function combinations, especially questions Expressions about psychological and emotional state i.e. doubt, tiredness and task difficulty
	Passive/active	The use of language functions to appear to solve a problem	<b>Semi-recursive</b> Limited range of language functions used Absence of doubt, tiredness and task difficulty
<b>Representational agency</b> Spoken	Productive	Selection, control and representation of personal meaning in relation to topic in which ideas and/or information is taken up by their peer	<b>Recursive</b> Shared information with ideas taken up Language functions relate to experience, (dis)interest, likes and desires in relation to topic
	Passive/active	Selection, control and representation of personal meaning in relation to topic	<b>Semi-recursive</b> Individualistic delivery Language functions relate to learner experience, (dis)interest, likes and desires in relation to topic

important for SLA and for the realisation of learners' assumed life goal of learning a language. The role of new types of agency and their relationship with language learning is yet to be determined. Whereas interaction as strategic agency may sustain learners in extended interaction, interaction as representational agency may support other learning benefits such as task personalisation or enhancing motivation.

Furthermore, if network-based learning scenarios are characterised by an expanded range of semiotic resources across multiple sites and time frames, typical spoken interaction tasks may not fit. In addition, assessment of task performance as "process" rather than "outcome" (Skehan, 1996) appears an important focus for online tasks because interaction cannot be assumed to be recursive and spontaneous in the TL. The need for conceptual models, based on empirical work, is needed to support design processes and analysis of learning activity in digital spaces for further enhancements (Blin & Jalkanen, 2014).

The study highlights the potential transferability of the results to other tasks and contexts with similar characteristics as a future potential research area. Further research is needed to confirm if new interaction types (i.e. passive/active) are occurring in other online tasks. We hope that the framework presented may be useful in identifying the characteristics of interaction thought to lead to optimal gains in SLA – that is to say shared, negotiated and recursive.

## References

- Ahearn, L. M. (2001) Language and agency. *Annual Review of Anthropology*, **30**: 109–137.
- Anderson, B. (1983) *Imagined communities: Reflections on the origin and spread of nationalism*. New York: Verso.
- Appel, C., Nic Giolla Mhichíl, M., Jager, S. and Prizel-Kania, A. (2014) SpeakApps 2: Speaking practice in a foreign language through ICT tools. In: Jager, S., Bradley, L., Meima, E. J. and Thoušny, S. (eds.), *CALL Design: Principles and Practice*. Dublin: Research-publishing.net, 12–17. doi:10.14705/rpnet.2014.000187
- Appel, C., Robbins, J., Moré, J. and Mullen, T. (2012) Task and tool interface design for L2 speaking interaction online. In: Bradley, L. and Thoušny, S. (eds.), *CALL: Using, Learning, Knowing*. Dublin: Research-publishing.net, 15. doi:10.14705/rpnet.2012.9781908416032
- Blin, F. and Jalkanen, J. (2014) Designing for language learning: Agency and languaging in hybrid environments. *Apples – Journal of Applied Language Studies*, **8**: 147–170.
- Breen, M. (1987) Learner contributions to task design. *Language Learning Tasks*, **7**: 23–46.
- Coughlan, P. and Duff, P. A. (1994) Same task, different activities: Analysis of a SLA task from an activity theory perspective. In: Lantolf, J. and Appel, G. (eds.), *Vygotskian approaches to second language research*. Norwood, NJ: Ablex, 157–172.
- Council of Europe (2001) *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg: Language Policy Unit.
- de la Colina, A. A. and García Mayo, M. P. (2007) Attention to form across collaborative tasks by low-proficiency learners in an EFL setting. In: García Mayo, M. P. (ed.), *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 91–116.
- Deters, P., Gao, X., Miller, E. and Vitanova, G. (eds.) (2015) *Theorizing and analyzing agency in second language learning*. Clevedon: Multilingual Matters.
- Dewaele, J. (2009) Individual differences in second language acquisition. In: Ritchie, W. C. and Bhatia, T. K. (eds.), *The new handbook of second language acquisition*. Bingley, UK: Emerald Group Publishing, 623–646.
- Dooley, M. (2011) Divergent perceptions of telecollaborative language learning tasks: Task-as-workplan vs. task-as-process. *Language Learning & Technology*, **15**(2): 69–91.

- Dörnyei, Z. and Ushioda, E. (2009) Motivation, language identities and the L2 self: Future research directions. In: Dörnyei, Z. and Ushioda, E. (eds.), *Motivation, language identity and the L2 self*. Clevedon: Multilingual Matters, 350–356.
- Doughty, C. J. and Long, M. H. (2003) Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, 7(3): 50–80.
- Ellis, R. (2000) Task-based research and language pedagogy. *Language Teaching Research*, 4(3): 193–220.
- Foster, P. and Ohta, A. S. (2005) Negotiation for meaning and peer assistance in second language classrooms. *Applied linguistics*, 26(3): 402–430.
- García, O. (2009) Education, multilingualism and translanguaging in the 21st century. In Mohanty, A. K., Panday, M., Phillipson, R. and Skutnabb-Kangas, T. (eds.), *Multilingual education for social justice: Globalising the local*. New Delhi: Orient BlackSwan, 140–158.
- Gass, S. M. and Mackey, A. (2006) Input, interaction and output: An overview. *AILA Review*, 19: 3–17.
- Halliday, M. A. K. and Hasan, R. (1989) *Language, context, and text*. London: Oxford University Press.
- Kumpulainen, K. and Mutanen, M. (1999) The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction*, 9(5): 449–473.
- Lafford, P. A. and Lafford, B. A. (2005) CMC technologies for teaching foreign languages: What's on the horizon? *CALICO Journal*, 22(3): 679–709.
- Lai, C. and Li, G. (2011) Technology and task-based language teaching: A critical review. *CALICO Journal*, 28(2): 498–521.
- Lamy, M.-N. (2006) Interactive task design: Metachat and the whole learner. In: García Mayo, P. (ed.), *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 242–264.
- Lantolf, J. (ed.) (2000) *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Luckin, R. (2010) *Re-designing learning contexts: Technology-rich, learner-centred ecologies*. Abingdon: Routledge.
- Martin, J. (2004) Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, 39(2): 135–145.
- Mercer, N. (2010) The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology*, 80(1): 1–14.
- Mercer, S. (2011) Understanding learner agency as a complex dynamic system. *System*, 39(4): 427–436.
- Miller, E. (2012, Sep 2) Chapter proposals: Agency and SLA. <http://linguistlist.org/issues/23/23-3666.html>
- Novick, D. G. and Sutton, S. (1997) What is mixed-initiative interaction? In Haller, S. and McRoy, S. (eds.), *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*. AAAI/MIT Press, 114–116.
- Ortega, L. (1997) Processes and outcomes in networked classroom interaction: Defining the research agenda for L2 computer-assisted classroom discussion. *Language Learning & Technology*, 1(1): 82–93.
- Pavlenko, A. and Lantolf, J. (2000) Second language learning as participation and the (re)construction of selves. In: Lantolf, J. (ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press, 155–177.
- Reeve, J. and Tseng, C. (2011) Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4): 257–267.
- Roebuck, R. (2000) Subjects speak out: How learners position themselves in a psycholinguistic task. In Lantolf, J. P. (ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press, 79–95.
- Samuda, V. and Bygate, M. (2008) *Tasks in second language learning*. Basingstoke: Palgrave Macmillan.

- Schwartz, D. and Okita, S. (2009) The productive agency in learning by teaching. Unpublished manuscript. [http://aaalab.stanford.edu/papers/Productive\\_Agency\\_in\\_Learning\\_by\\_Teaching.pdf](http://aaalab.stanford.edu/papers/Productive_Agency_in_Learning_by_Teaching.pdf)
- Skehan, P. (1996) Second language acquisition research and task-based instruction. In: Willis, J. and Willis, D. (eds.), *Challenge and change in language teaching*. Oxford: Heinemann, 17–30.
- Swain, M. (2006) Linguaging, agency and collaboration in advanced second language proficiency. In: Byrnes, H. (ed.), *Advanced language learning: The contribution of Halliday and Vygotsky*. London: Continuum, 95–108.
- Swain, M. and Deters, P. (2007) ‘New’ mainstream SLA theory: Expanded and enriched. *The Modern Language Journal*, **91**(s1): 820–836.
- van Lier, L. (2008) Agency in the classroom. In: Lantolf, J. and Poehner, M. (eds.), *Sociocultural theory and the teaching of second languages*. London: Equinox, 163–186.
- Wang, J. (1996) Same task: different activities. Unpublished research report. Pittsburgh, PA: University of Pittsburgh.
- Xiao, J. (2014) Learner agency in language learning: The story of a distance learner of EFL in China. *Distance Education*, **35**(1): 4–17.
- Yim, Y. K. K. (2011) Second language students’ discourse socialization in academic online communities. *Canadian Modern Language Review*, **67**(1): 1–27.
- Zhang, J. (2010) Learner agency, motive and self-regulated learning in an online ESL writing Class. *IALLT Journal*, **43**(2): 57–81.

17.2 Paper 2: The Negotiation of Shared and Personal meaning making  
in Spoken Interaction Tasks



## The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks

Janine Knight

Universitat Oberta de Catalunya and Universitat Internacional de Catalunya

[jknight@uoc.edu](mailto:jknight@uoc.edu)

[janine@uic.es](mailto:janine@uic.es)

Elena Barberà

Universitat Oberta de Catalunya

[ebarbera@uoc.edu](mailto:ebarbera@uoc.edu)

*This paper reports on learner foci for jointly-constructed and individual meaning making in three typical spoken interaction tasks. It explores negotiation processes including personal meaning making whilst 'languaging'. Three tasks were analysed: an information gap, opinion sharing and role-play task to ascertain meanings generated by learners as they (re)configured tasks with the semiotic resources available. Audio-transcripts from peer-to-peer synchronous, Computer Mediated Communication (SCMC) tasks were analysed using content analysis for studying learner purposes and an emic perspective for studying learners' mediation of semiotic resources. Results were triangulated with other data sources.*

*Results suggest that the foci for meaning negotiation through speech across tasks spanned four areas, namely 1) task management 2) completion of pedagogical task 3) relating self to task topic and 4) co-ordinating technological task features. Online tasks require a greater need for dyads to negotiate organisational moves as well as the need to attend to technological aspects, including difficulties. Different task types channelled different task completion strategies and regulated the ability to personalise the topic.*

*Findings reveal how learners' semiotic budgets can be expanded in relation to technological aspects of tasks, which learners act upon and which can shape how and what is negotiated, offering new opportunities for meaning making. Pedagogical implications for task design are proposed.*

**Keywords:** Language learning, meaning negotiation, spoken interaction, languaging, task type, design

### 1. Introduction

The ubiquity of technology in everyday life as well as the many digital environments inhabited for work, play or socialisation, offers an ever expanding 'semiotic budget' thus providing increased opportunities for languaging ...as we engage in diverse activities [1]. 'Languaging' is an action or process involving "making meaning and shaping knowledge and experience through language" [2] (p.98). In task design, semiotic resources can include images and/or text offering visual and/or audio 'inputs' [3] such as instructions and/or image. Another realm where semiotic resources can reside is within technological tools that can host tasks. Resources can be mediated through learners' physical actions, visual abilities and language during task processes. 'Inputs' e.g. task instructions and an image have typically assumed a fixed, constant learning space and synchronous time mode (i.e. face-to-face classroom). However, in online tasks the semiotic budget is potentially expanded in terms of number of resources residing in different realms and the fact that learners can move across and within different realms. This raises questions as to whether tasks fit digital contexts. This study explores this potential mismatch by focusing on how and for what purposes learners make meaning in three online tasks in order to inform future online task design.

### 2. Theoretical Framework

A sociocultural perspective encompasses the notion of 'inputs' as 'semiotic resources' or 'tools' [4] as a way of understanding how learning is mediated through the use of printed materials, physical environment, gestures and classroom discourse [5]. Tools (or semiotic resources) can be a physical artefact ... or symbolic, as in the case of utterances produced during conversations with others and



the self [6]. The environment “is full of potential meanings [...that] become available gradually as the learner acts and interacts within and with this environment” [7] (p.246).

In online tasks, semiotic resources can form part of ‘task-as-workplan’ [8]: resources that task designers intend users to mediate with e.g. a virtual button on an interface, textual instructions and/or linguistic input from texts or audio/visual images. Other resources may also be employed by learners during ‘task-as-process’ [9]: resources actually employed. In the mediation process, where meaning making takes place, the analysis of thinking and speaking is not focused on semiotic resources in the task-as-workplan, but on “the active learner, or activity itself” [10] in ‘tool-mediated goal-directed action’ [11]. However, task ‘goals’ have been conceptualised differently: 1) as task completion [12]; 2) the goal of activity/ies during tasks [13]; 3) learners own goals brought to the task [14] and 4) as emerging during processes according to the needs and goals of the moment [15] (p79). In online speaking tasks goals can be achieved through mediational means using any number and/or combination of semiotic resources. These include intentionally using the language system or ‘speech as action’ [16] alongside ‘sensory and motor systems’ [17] to carry out ‘agentic actions’ [18].

The use of pedagogical tasks has migrated from face-to-face to online contexts. Tasks can be seen as ‘a cognitive device’ [19] to elicit spoken interaction or meaning negotiation between peers. The term ‘Negotiation for Meaning’ (NfM) has been conceptualised differently within cognitivist and sociocultural perspectives with different emphasis given depending on “particular theoretical orientations of...writers” [20] (p. 74). Although cognitivist perspectives include using it to refer to non-understanding sequences (from Varonis and Gass, 1985) [21], Block [22] proposes that the term has omitted other interpretations such as negotiation of solidarity and support [23] and face [24]. This study conceptualises meaning negotiation as ‘language-as-action’ (languaging): exchanging information between peers (or negotiation *for* meaning) *and* for different purposes (or negotiation *of* meaning) because “what is achieved by negotiation varies according to the purpose of negotiation” [25] (p 74). Meaning can also be made through ‘conversations with the self’ [26].

### 3. General objectives

This study explores how learners act with semiotic resources to negotiate meaning in order to achieve smaller goals by focusing on tool-mediated goal-directed action across three tasks. The research question is: *What are the foci of learners’ goal-directed actions and how do they as dyads and individuals negotiate meaning of content?*

### 4. Method

Participants were adult learners enrolled on English as a Foreign Language class as part of their degree programme at the *Universitat Oberta de Catalunya*, a 100% online university based in Barcelona. All learners were considered bilingual, sharing Catalan and Spanish. Eight participants (four dyads) from the first data set (completing an Information gap and Opinion sharing task) had a level of B2.1 on the Common European Framework of Reference for Languages whereas participants from the second data set (completing a role-play) had a level B1.1. The technological tool used for synchronous spoken interaction was the Tandem tool <http://www.speakapps.eu/#tandem>. The first data set pertained to an exploratory study focusing on learner agency [see 27]. The current study combines two data sets in order to focus on meaning negotiation across different tasks. The exploratory study also revealed (through student questionnaires) that cases 2, 3 and 4 looked at answers before recording interaction, impacting on some findings in this study.

Regarding instruments, language functions from Kumpulainen and Wray (2002) [28] were used to code turns using content analysis. Speech units were identified on an utterance basis. The unit of analysis was each dyad and individuals within that dyad. A purposive sampling approach was used to include pairs that demonstrated ‘good’ performance and those who approached the task differently to task design. Data review and analysis was carried out concurrently with data collection.

New language functions, when identified, were added to the original coding system iteratively and mapped across two data sets. New codes were checked for inter-rater reliability. The emic (insider) perspective was taken to examine subjective experiences of participants alongside an interpretative approach to collection and analysis. Activity logs, student guidelines, screenshots of researcher-conducted task simulations and teacher checks triangulated findings.





## 5. Results and discussion

The analysis of learner talk alongside an identification of what learners were looking at and/or touching revealed how learners shifted their focus and purposes depending on goals moment-by-moment. This spanned across four general areas, namely: 1) task management; 2) pedagogical task completion; 3) talking about individual self in relation to task topic and 4) co-ordinating navigational aspects. Learners made meaning both jointly and/or individually relating to these four foci.

### Task Management

Language was the main resource for management of organisational purposes but technological task features and other tools were also used, including the Tandem interface pages. Organisational talk occurred when the need for management emerged: beginning of a task; transition between tasks; end of a task as well as when learners had to manage unpredictable technological problems that 'disrupted' the pedagogical task, such as headphones falling (case 1), being shown a 'time is up' pop-up while partner was on the wrong page (case 5), or ascertaining whether the recorder had worked (case 7). The need for learners to manage pedagogical and technological task features highlights that opportunities for making meaning are expanded in online tasks. Explanations may include teacher absence to organise learners, learners needing to 'negotiate the unexpected' technological problems. Self-talk was also a feature of this focus.

### Pedagogical Task Completion

This focus elicited the use of language functions for task completion, which varied across task types and dyads. The varied combination of language functions by individuals and dyads highlighted different strategies for making meaning. Although language was the main semiotic tool used for task completion, other resources were used: photo, textual instructions and prompts (Information gap and opinion sharing), textual instructions and prompts (role-play task). Cases 2, 3 and 4 prepared their interaction and knew an uncommon word 'awnings', implying that they had used a virtual or paper-based dictionary to find the meaning. In addition, sounds of paper rustling in recordings and reference to "the pdf" from case 7 suggested that apart from official pedagogical 'inputs' and technological features, learners employed other 'non-official' resources.

### Talking about Individual self in relation to Task Topic

This focus was achieved through the use of a photo of a London street and textual instructions of an open question in the opinion sharing task. In the role-play task, this was achieved through role allocation. The interviewee gave personal information, related experiences, and preferences in response to a series of questions indicated for the interviewer through textual instructions and prompts. Talking about the self therefore occurred simultaneously to pedagogical task completion through questions/answer sequences. The Information gap task did not support personal meaning making in relation to topic.

### Co-ordinating Navigational Aspects

The Tandem tool required both learners to navigate physically through touch and orally through speech using virtual buttons to start, submit answers and go to different task pages. Because learners could not see each other or each other's screens they had to negotiate intentional moves: what they wanted to do and when, co-ordinating navigational aspects. The icons and buttons of the Tandem became resources with which to negotiate meaning around, becoming 'inputs', the focus of interaction. However, the extent of which this occurred depended on if and how dyads followed the 'official' route/sequence in the 'task-as-workplan'. Many icons and buttons were notably (consistently) absent in some transcripts indicating dyads had partially followed the intended navigational route to the task instructions but afterwards created their own 'lines of desire' (an architectural term referring to paths people make, which are often shortcuts that ignore the given route) [29].

## 6. Conclusions

Online tasks appear to expand the realms and foci for shared meaning negotiation particularly for organisational purposes including unexpected events and negotiating navigational paths. Because smaller goals emerge during task processes according to needs and goals of the moment [30] pedagogical 'task completion' is only one making as learners multi-task. Future research might explore challenges posed by an expanded semiotic budget whereby (pedagogical) input and technological features of host tools are only a few of the potentially multiple foci for meaning negotiation.





The different foci highlights the need to linguistically scaffold learners to deal successfully with multi-tasking such as making specific words/phrases related to technological aspects available (e.g. “timer”, “headphones”), particularly if supporting learners maintaining extended recursive interaction in the TL is an aim.

## References

- [1] Blin, F., & Jalkanen, J. (2012). Agency and languaging: Exploring design-based pedagogies for language learning. Paper presented at the *Insights into Applied Linguistics: Languaging, Agency, and Ecologies Conference, Jyvaskyla, Finland*,
- [2] Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (ed.), *Advanced Language Learning: The contribution of Halliday and Vygotsky*. London: Continuum, 95-108.
- [3][19][25] Ellis, R. (2000). Task-based research and language pedagogy. *Language teaching research*, 4(3), 193-220.
- [4] Vygotsky, L. S. (1981). The instrumental method in psychology. *The concept of activity in Soviet psychology*, 135-143.
- [5][10] Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford University Press.
- [6][26] Wertsch, J. V. (1998). *Mind as action*. Oxford: Oxford University Press.
- [7] Van Lier, L. (2000). 11 From input to affordance: Social-interactive learning from an ecological perspective. *Sociocultural theory and second language learning*, 245.
- [8][9] Breen, M. (1987). Learner contributions to task design. *Language Learning Tasks*, 7, 23-46.
- [11] Zinchenko, V. P. (1985). Vygotsky's ideas about units for the analysis of mind. *Culture, communication and cognition: Vygotskian perspectives*, 94-118.
- [12] Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 17-30). Oxford, UK: Heinemann.
- [13] Wang, J. (1996). Same task: different activities. *Unpublished research report. Pittsburg, PA: University of Pittsburg*.
- [14] Coughlan, P., & Duff, P. A. (1994). Same task, different activities: Analysis of a SLA task from an Activity Theory perspective. *Vygotskian approaches to second language research*, 173-193.
- [15][30] Roebuck, R. (2000). Subjects speak out: How learners position themselves in a psycholinguistic task. *Sociocultural theory and second language learning*, 79-95.
- from an activity theory perspective. *Vygotskian approaches to second language research*, 173-193.
- [16] Austin, J.L. (1962). *How to do things with words*. Oxford: Oxford University Press.
- [17][18] Bandura, A. (1999). A social cognitive theory of personality. In L. Pervin & O. John (Ed.), *Handbook of personality* (2nd ed., pp. 154-196). New York: Guilford Publications.
- [20] Ellis, R. (2001). The metaphorical constructions of second language learners in Breen (ed.) *Learner contributions to language learning: New directions in research*, London. Longman 65-85.
- [21] Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied linguistics*, 6(1), 71-90.
- [22] Block, D. (2003). *The Social Turn in Second Language Acquisition*. Georgetown University Press, 3240 Prospect Street, NW, Washington, DC 20007.
- [23] Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied linguistics*, 19(1), 1-23.
- [24] Davies, J. (2000). *Social Perspectives on negotiation for meaning*, unpublished MA dissertation, Institute of Education, University of London
- [27] Knight, J. Barbera, E. & Appel, C, (2016). *Learner Agency in Online Spoken Interaction tasks*. Manuscript submitted for publication
- [28] Kumpulainen, K., & Wray, D. (2002). *Classroom interaction and social learning: From theory to practice*. Psychology Press.
- [29] Luckin, R., & Du Boulay, B. (2003). Lines of desire: the challenges of interactive educational TV. In *11th international conference on artificial intelligence in education (AIED2003)*. Sydney, Australia.

17.3 Paper 4: Navigating a multimodal ensemble: mediating turns verbally and non-verbally in online interaction

## **Navigating a multimodal ensemble: learners mediating verbally and non-verbally in online interaction**

Research into the multimodal aspects of language is increasingly important as communication through a screen plays a greater role in modern society than ever before (Liou, 2011) e.g. computer screen, mobile or tablet. Multimodality has been explored from a number of angles relating to Computer Mediated Communication (CMC), such as its affordances and impact on language learners, highlighting its relevance and importance in the field of second language acquisition (SLA). Because attendance to both peers and the screen is required in CMC scenarios, learners can be seen as positioned as 'semiotic initiators and responders' (Coffin and Donohue, 2014). Increasingly researchers are highlighting a need for a methodological 'turn' to analyse this scenario from a 'language' focus to a more holistic understanding of the interactions (Kress and van Leeuwen, 2001; Lamy, 2006; Hampel and Hauck, 2006; Flewitt, 2008).

Along these lines, a case study approach is used to explore how the action of task completion is mediated between 6 dyads (and individuals within the dyads) during an online peer-to-peer spoken interaction audioconferencing event. Drawing on notions from multimodal (inter)actional analysis (Norris, 2004; 2006) and the notion of 'semiotic initiators and responders', semiotic mediation - with screen-based resources- is explored. Different data sources and methods are employed including audio recordings, screenshots, log files, task simulation and reconstruction.

Results highlight the semiotic initiations and responses-including oral turn-taking in task completion presented as a framework.

**Key words:** task-based synchronous communication (TB-SCMC), spoken interaction, semiotic mediation, multimodal turn-taking, screen-based resources, agency

### **1. Introduction**

The many digital environments inhabited for work, play or socialisation, offer an ever expanding semiotic budget (Blin and Jalkanen, 2014) for learners in online language learning programmes. A semiotic budget may include the potential visual, textual and aural inputs that learners use as resources in interaction for task completion. Analysis of how these semiotic resources form part of communication and representation in Computer-mediated communication (henceforth CMC) can be understood through the lens of multimodality that conceives communication as much more than just spoken or written language.

Understanding online language learning through a multimodal lens is important for a number of reasons. Visual elements through a screen, particularly multimodal texts, are gaining prominence for communication in modern society (Liou, 2011). There is a need for studying CMC task-based events through this lens because current language learning technologies incorporate more graphic, visual, textual and auditory information (Collentine, 2009), often facilitated by 'media convergence' (Herring, 2015) of modes in digital scenarios. In addition, multimodality should be taken into when it comes to task design (Canto, de Graff and Jauregi, 2014; Hampel, 2010; Hauck, 2010).

In tasks facilitated through audioconferencing, learners may be exposed to a 'multimodal

ensemble' (Bezemer and Kress, 2016) which consists of such screen-based resources requiring learners to be 'semiotic initiators and responders' (Coffin and Donohue, 2014). Learners mediate this ensemble in many different ways –at times innovatively and differently from the initial design purpose.

However, there is a lack of research on the impact of multimodal communication in online language classrooms (Hampel and Stickler, 2012) and multimodality in task-based classrooms in general (Gilabert, Manchón, and Vasylets, 2016). Methodological approaches to analysis are still at an exploratory stage (Rossolatos, 2015) and remain a challenge (Herring, 2015) despite a few key studies (see Flewitt, Hampel, Hauck and Lancaster, 2017 and ReCALL Special Issue, September 2016). This study aims to help fill that gap by exploring how the (semiotic) mediation for task completion is carried out whereby learners may be positioned as 'semiotic initiators and responders' (Coffin and Donohue, 2014).

## **2. Theoretical framework: Semiotic mediation and CMC in language learning**

The terms 'semiotic resources' and 'tools', introduced by Vygotsky (1981) and highlighted by Van Lier (2000) and Lantolf (2000) within a sociocultural perspective of language learning is a way of understanding how language learning is mediated through different available resources that students encounter during online tasks. During task completion, semiotic resources can mediate goal-directed actions. Mediated action involves agents and their cultural tools –both are mediators of the action (Wertsch, 1998). Language acquisition occurs through a dynamic process of the self interacting with cultural tools; this 'mediation' includes the use of language as a social resource, which becomes internalised by the learner. This cognitive development occurs moment by moment in social interaction (Lantolf, 2000). The microanalysis of discourse in its sequential context is therefore important because it allows the researcher to examine this process in motion (Lantolf, 2000).

An appreciation of how cultural tools or mediational means are involved in human action (including learning) forces us to go beyond the individual agent when trying to understand the forces that shape such action (Wertsch, 1998). Therefore, attending to "the material *stuff*" (Kress, 2003, p. 32, original emphasis echoed by Lamy, 2006), in this case the materiality of the screen, suggests that any analysis of mediated action for the completion of a spoken interaction task should consider multiple ways of understanding the action.

In online CMC scenarios, trying to understand these complex situations that involve mediation through human-to-human oral interaction as well as mediation that takes place as a result of human-to-computer-interaction requires expanding the analysis to include non-embodied screen-based semiotic resources. By non-embodied modes we mean screen-based resources that may not form part of the direct two-way *communication* but rather are *represented* from computer to learner. To exemplify, in CMC facilitated by audioconferencing, audio is a mode emanating from the oral utterances from their partner *through* the channel but it is also possible for audio to be a mode represented by the screen *with* which learners can also mediate (e.g the sound from videoclips).

Learners can be both 'semiotic initiators' and 'semiotic responders' (Coffin and Donohue, 2014): initiating and responding orally with their peer through oral turns and

turn-taking as well as to various screen-based resources - a notion we subsequently return to.

There are increasingly more CMC studies for spoken interaction that highlight the screen-based modes and interface. Hampel and Stickler (2012), in a videoconferencing event, identified communication modes as linguistic (spoken and written) alongside visual such as icons (vote buttons yes/no/? emoticons), still and moving images, display/scrolling of text and gestures. Lamy's (2006) study on turn-taking and facesaving using an audiographic tool identified natural language (written and spoken) as well as visual resources (icons, images, colours and shapes). Vetter and Chanier's (2006) study on how language learners used multimodal tools to make spoken interactions highlighted text, speech, graphics for communication and highlighted the interplay of modes (text and spoken language). Knight and Barbera's (2016) study of peer-to-peer spoken interaction tasks using an audioconferencing tool found that learners were multi-tasking as they interacted with language (text), image (photo), icon (pop-up) and navigational buttons. The different screen-based resources appeared to relate to different learner purposes. Balaman and Sert's (2017a) study on two different task types in two different settings (face-to-face and online using audioconferencing) highlighted the videoclips on the screen whereby learners could type answers, click on answer buttons and receive correct answers whilst conversing with their partner. Whereas in the face-to-face task learners orientated to topic maintenance, in online tasks orientation was to the interface and on screen activities.

In another study, Balaman and Sert (2017b) confirmed how participants coordinated their (oral) interactions with their orientations to the task interface. Learners' co-ordination of multiple actions tasks was highlighted in Knight and Barbera's aforementioned study (2016).

## **2.1 Semiotic Initiation and Response and multimodal turn-taking**

'Semiotic initiation' and 'response' (Coffin and Donohue, 2014) in CMC tasks in the field of SLA has largely been studied through the analysis of verbal turn-taking, often using conversational analysis (henceforth CA). Whereas CA's originally purely verbal outline of the turn-taking 'system' (sometimes referred to as the speech exchange system) as proposed by Sachs, Schegloff and Jefferson (1974) has predominated, there has been a growing interest in the multi-modal dynamics of the turn-taking process in the field of linguistics, as well as in SLA (see overview by Jenks, 2014). More recent studies (some beyond the field of SLA) have expanded oral turn-taking to include modes emanating from humans including gaze behaviour (e.g. Oben and Brône, 2015) with others highlighting the need to take gesture into account (e.g. Mondada, 2007; 2013).

Some CA analysts, both in online language learning and non-language learning environments, found that turn-taking can take place transmodally - across modes (Lamy, 2006 in audiographic conferencing; Liddicoat, 2010 and Helm and Dooly, 2017 in videoconferencing). Both Lamy (2006) and Helm and Dooly (2017) highlight the use of hybrid, mixed mode interaction (with text and speech), often times resulting in time lags and overlapping of turns. Key to this understanding of online, multimodal turn-taking is the way in which speakers in each oral turn demonstrated to one another their own understanding of the previous speakers' oral turn and that these aspects of turns and turn-taking were context-sensitive to both task type and task setting (cf. Balaman and Sert, 2017a). Long silences suggested learners were orientating to the screen. These authors

show that learners coordinate turn-taking through their mutual alignment to orientations to screen-based resources, online oral interaction, and other features of the interface.

Whilst CA analysts maintain that CMC conversations can involve a multimodal accomplishment of openings, interruptions and closings as oral turns carried out in various mediums (Tudini, 2014), Liddicoat (2010) highlights an additional further complexity: it can be argued that turn-taking can take place between humans and screen-based resources not just between humans. In his study, Liddicoat (2010) found that the initiator in the beginning of an online conversation must firstly secure a non-present co-participant which must be achieved through the technology. This was achieved by a message via the computer (screen) namely 'Andrew wants to have a video conversation' which is neither spoken or written by Andrew but is initiated by him with the technology. The response of his partner is either a choice to press 'respond' or 'refuse'. This resembled summons-answers sequences where the verbal equivalent may be 'hey' or naming and the technological equivalent may be the ringing of the telephone (Liddicoat, 2010). The online nature of the participants' interaction was considered a relevant constituent part of the interactions (turns) not just a facilitative one (Liddicoat, 2010).

Beyond CA studies, turn-taking and turns with computers has been extensively studied in the field of Human Computer Interaction (henceforth HCI) and CMC which this paper cannot do justice to due to word limitations. However, pertinent (non SLA) studies have highlighted that screen-based resources can be 'active agents' that send reminders (Dourish and Belotti, 1993) as turns, agents in conversation characterised by 'presentation' and 'acceptance' phases (Clark and Brennan, 1991) as well as be used as conversational resources in the accomplishment of physical tasks (Kraut, Fussell and Siegel, 2003).

More recently, Benson (2015) from a digital discourse perspective considered physical turn-taking with the interface of YouTube and employed the notion of 'orientation' from CA to understand turns. He operationalised 'responses' on a YouTube page as video responses, 'like'/'dislike' icons or written comments. Turn 'initiation' included uploading a video -highlighting the visual, textual and physical modes involved. What these HCI studies highlight is that screen-based resources can appear to act out turns (as initiators or responders) and that physical moves of human participants in relation to them may also be characterised in this way.

Taking this into account, the screen-based resources can be understood through the notion of potential 'others' that may also act as and/or be orientated to in this way (Raudaskoski, 1999). The study of 'others' has generally been approached as being 1) text or discourse; 2) a social entity or agent or 3) a sign (Raudaskoski, 1999). Human participants may also be positioned as the 'other' in different 'encounters' rather than 'conversations' meaning that in pairwork screen-based resources may act as others (Raudaskoski, 1999). However, because the semiotic resources are not "fully fledged communicative partners..." "the sense making has to be constructed one-sidedly, rather than co-constructed, making the human participant solely responsible for the emerging meaning" (Raudaskoski, 1999, p. 22-23).

This study conceptualises oral turns and turn-taking within the notion of 'semiotic initiation and response' (Coffin and Donohue, 2014) because we deem it to be a more holistic notion considering the different possible forms that screen-related 'turns' may

take.

### **3. Purpose of study and research questions**

The purpose of this study is to understand how learners carry out the mediated action of completing an online speaking task. With the studies presented and the gaps in literature in mind, this study aims to create a framework for understanding the interactions taking place *with* the screen-based resources and *through* the screen, whereby orientation to potential screen-based ‘others’ (Radauskoski, 1999) may be more fully explored. We also want to know whether discourse descriptions of ‘semiotic initiations and responses’ including oral turn-taking are useful notions to characterise mediation taking place between humans and the screen not only human-human interaction. Therefore, the research question is:

How do learners carry out mediated action as task completion through and/or with the screen-based resources in an online spoken interaction task?

### **4. Analytical approach and methods for analysis**

A qualitative case study approach was employed incorporating a purposive sampling procedure in order to select dyads that had appeared to follow task design and others that appeared to diverge from it. Data sources were triangulated with confirmation checks with colleagues about the tasks given and task conditions. Researcher notes were made and tables were constructed in order to compare various similarities and differences between the cases and individual learners’ behaviour.

For our analytical perspective to the data, we operationalized mediated action by drawing on notions from multimodal (inter)actional analysis developed by Norris (2004) and notions from conversation analysis (henceforth CA) developed by Sachs, Schegloff and Jefferson (1974), in particular the orchestration of turn-taking. However, this study does not rigorously follow the protocols of CA transcription since the focus is on the non-oral sequentiality of initiations and responses including oral turn taking.

We also draw on the principles of ‘relevance’ and ‘orientation’ from CA which Benson (2015) used in his identification of participants physical turns with the screen. We operationalise learners’ physical turns- which can be initiations or responses- with the screen as when learners indicate clicking a screen-based resource. In addition, we operationalise screen-based turns or initiations as when learners respond to them orally, physically or visually (or a combination of both). In this way our approach uses an emic (learner) perspective and etic (researcher) perspective.

When learners explicitly mentioned a resource orally we deemed this ‘relevant’ considering that “modes do not exist without social actors utilising them in some way” (Jewitt, Bezemer and O’Halloran, 2016 p. 115). We looked for instances of ‘semiotic initiations’ and ‘responses’ (Coffin and Donohue, 2014) (that also encompassed oral turn-taking) from both learner and screen. We focused on learner’s oral turns as responses to previous oral turns and learner’s potential response/initiations to the screen-based resources. We identified what a turn’s talk was occupied with or the learner’s

understanding of the previous turn/initiation and if and how learners “react to the messages” (Norris, 2006 p. 4).

We refer here to our complementary use of multimodal (inter)actional analysis. This analytical approach identifies three levels of social (mediated) action: higher-level, lower-level actions and frozen action, each of which deals with a different level of interaction. *Higher-level action* is used to refer to large-scale actions, such as a meeting, and is made up of ‘multiplicity of chained lower-level actions’ (Norris, 2004). *Lower-level action* is used to refer to smaller-scale actions, for example, gestures or gaze shifts that become chains of lower-level interactions (Norris, 2004). The lower-level actions support the achievement of the higher-level action. Higher-level mediated actions are those actions that social actors usually intend to perform and/or are aware of and/or pay attention to (Norris, 2016). *Frozen actions* are entailed in material objects after the action has taken place such as the layout of a room. However, because we cannot observe the physical space surrounding the learners’ computers, this current study does not apply the notion.

Multimodal (inter)actional analysis also deploys the notion of levels of simultaneous awareness/attention namely *foreground*, *mid-ground* and *background*. This is a continuum that facilitates the visual representation of various levels of attention that an individual is simultaneously engaged in (Norris, 2016) whilst completing an action. A person can be engaged in various actions at a particular point in time (e.g. engaging in a research project, Skyping with family members, interacting with girlfriend) (Norris, 2016) and so decreasing attention/awareness can be described by the continuum.

## 5. Instruments and Analysis Procedure

The unit of analysis is ‘mediated action’ (Wertsch, 1998). In order to be familiar with ‘the multimodal ensemble’ (Bezemer and Kress, 2016) presented on the screen and how the task design ‘guided’ a learner in task completion, the researcher simulated the task process synchronously with a colleague, taking screenshots to document the task process (see Appendix A). Based on this, a labelling process of the screen-based resources (and what appeared on the screen in-between) was made (see Appendix B). These were then categorized following Lamy (2006) who identified natural language in its written and spoken forms, as well as visual resources such as icons and images. In our categorisation we add ‘navigational’ to ‘textual’ and ‘visual’ as well as the terms ‘static’ and ‘dynamic’ to indicate whether screen-based resources are moving or not, the latter two labels taken from Herring’s (2015) rubric of multimodal CMC.

A key term employed in our analysis is resource (e.g. a visual timer, pop-up on the screen). Where a number of modes make up a resource we name the resource according to the sequence of modes that learners are intended to mediate with or the hierarchy of the modes important to its designed purpose. So for example, a pop-up that appears suddenly, needs to be seen first, then read to know what to do with it, and then ‘closed’ with a physical click. This would be labelled as a visual/textual/navigational resource. Textual information pertaining to the task was labelled as textual/visual, whereas banners designed to signal task sequence were labelled visual/textual. However, we recognise this is problematic depending on what mode learners are attending to at any given moment. Appendix C illustrates the analysis of the interface pages and the labelling according to this logic.



Different foci are applied for the analysis of mediation: 1) jointly-constructed mediated action of the dyad and; 2) individually mediated action of the learner/interviewer with the screen. Only the resources made explicit verbally in the recorded talk were considered relevant by learners. Notes were made iteratively, dyad after dyad that were then transformed into a comparative table, as a cross-case synthesis. This resulted in a chronological overview of the trajectories of screen-based resource-use made explicitly relevant by the pairs in their talk.

A parallel analysis focused on the mediated action of individual learners from an etic perspective. We sought to reconstruct the learner's 'steps' (who had the interviewer's role) as if they were following the task design. The main researcher, through listening to the recordings with the screenshots in hand, followed if and how the interviewer responded to the screen-based resources. In this analysis the interviewers did not make the screen-based resources verbally relevant as a topic but they made them relevant by responding to them in some way (whereby the sense making is constructed one-sidedly by students - Radauskoski, 1999). This was noted and tracked for each case with the results transposed to the table from the first analysis,

## **6. Participants and Context**

The participants were students in an English as a Foreign Language class as part of an online degree programme. The learners were in a B1.1 level on the CEFR (upper intermediate) group. There are 12 adult students: 2 male and 10 female, 26-55 years old, engaged in a virtual synchronous peer-peer oral role-play task. Students are bilingual (Catalan and Spanish) with English as an additional language. Participant names have been changed.

Learners in the study are presented with the interface pages from the Tandem audioconferencing tool: <http://www.speakapps.eu/#tandem>, which is a content management application that distributes task materials in real-time. The 12 students form 6 dyads as case studies. Data sources include approximately 97 minutes of peer-to-peer recorded oral conversations of which approximately 53 minutes that were transcribed in broad transcription (before learners changed roles and repeated the task). Screenshots of the task (by the researcher) and 50% of the Tandem tool logs (which indicate the date, time, number of entries and duration the tool has been open for use) are also used in the analysis. The other 50% were not retrievable while additional cases were added.

The task was a role-play task (divided into sub-tasks) in which learners took turns being the interviewer and interviewee: the first task requiring that one learner ask questions to his/her peer in their roles. The second sub-task required the interviewer to describe two jobs that he/she has available on the interface page with the aim that his/her peer decide which one they preferred. After the first two sub-tasks are completed (Task 1 and 2 in the screenshots) the sub-tasks are repeated but the peers change roles (Tasks 3 and 4 in the screenshots). The task was designed to be timed that indicated minutes and seconds remaining for learners to complete a task. This was followed by a pop-up that appeared when the pre-determined time was up (4 minutes for task 1 and 7 minutes for task 2). Only the first two tasks were analysed as this was deemed sufficient to answer the research question. The analysis stops at the point when learners swap roles and begin to

repeat the two sub-tasks. This means that one student is analysed as interviewer not both students.

## **7. Analysis and Discussion**

Learners mediated with the various navigational, textual, visual screen-based resources as well as with spoken language in the form of initiation /response sequences for task completion. Specifically, oral turn-taking, initiation and responses relating to navigational resources and responses to pedagogical task instructions were found to be the three main strands of the mediated action.

### **Oral turn-taking**

Regarding spoken language, examples 1 and 2 show how learners typically used verbal turn taking to mediate the process of pedagogical task completion.

Example 1 Question/answer (Case 2):

L: Erm Well, (paper rustling) Thank you in advance for your time. I make you some questions about your Curriculum Vitae. First: what your complete name, please?

P: My name is Paulo Martinez.

L: Um What are your academic experience?

P: I have a degree of Psychology

Some of the oral turn taking revealed learners' orientations to the screen-based resources such as example 2, case 3 below so that even when technological problems arose during the role-play interview and learners had to 'close' or minimize a resource so they could continue, oral turn-taking through questions and answers, was maintained. Learners orient to the interface ('close' icon) as found by Balaman and Sert (2017a) and use the the text 'close' as a conversational resource in the accomplishment of the physical task as found by Kraut, Fussell and Siegel (2003).

Example 2 Question/answer (Case 3):

A: Okay... well ok. now... close, no, I suppose...?

F: I suppose too, I close it...

### **Screen initiations and screen or human responses**

Navigational and/or textual and/or visual resources were also found to mediate action through screen initiations and/or screen or human responses. In order to understand this process, a table was created to indicate different trajectories of learner-use of the screen-based resources. We focused specifically on when learners explicitly mentioned a resource then compared across the cases. These were labelled 'emic' in table 1. We return to the resources labelled 'etic' in due course.

Table 1. Screen-based resources that are and are not made verbally explicitly by cases

Key:

Visual/textual (dynamic) (e.g. Pop-up requiring no response)

Textual/visual (static) (e.g. Information and Instructions)

Visual/Textual/Navigational (dynamic) (e.g. Pop-up requiring response)

Visual/Textual/Navigational (static) (e.g. virtual Button)

Visual (dynamic)

IR = interviewer; IE = interviewee

EM=emic; ET= etic

	Waiting for confirmation T1	Information about roles- Student A/B T1	(1) Ask the minimum of 5 questions T1	Textual sample questions e.g. 'Where do you live?' T1	(2) Create own questions + symbol? T1	Text from sample candidate T1	(3) Pop-up 'Time up!'/ 'close' button	(4) Next task	Uses text describing two jobs T2	(5) Pop-up 'Time up!' / 'close' button	Interface page 2 'Solution's page'	Timer	(6) Next task
	EM	ET	ET	ET	ET	ET	EM	EM	ET	EM	EM	EM	EM
Case 1	X	X	√ Inter	√ Inter	√ Inter	√ Inter	X	√ Inter	√ Inter	√ Intee	X	X	√ Intee
Case 2	X	X	√ Inter	X	√ Inter	√ Inter	X	X	√ Inter	X	X	X	X
Case 3	√ (before start T1) Inter	√ Inter Intee	√ Inter	√ Inter	√ Inter	√ Inter	√ Inter Intee	√ Inter	√(7) Inter	√ Intee	√ Inter	X	X
Case 4	X	√ T1 Inter Intee	√ Inter	√ Inter	√ Inter	√ Inter	X	X	√ Inter	X	X	X	X
Case 5	X	X	√ Inter	√ Inter	√ Inter	√ Inter	X	X	√ Inter	X	X	X	X
Case 6	X	X	√ Inter	√ Inter	√ Inter	√ Inter	X	X	√ Inter	X	X	√(8) Inter	X

(1), (2), (3), (4), (5) and (6) considered to be a turn in the form of a request or invite that learners can respond or accept in order to complete the turn.

(7) Case 3 refers to "the pdf" after tool does not work.

(8) Case 6 refers to the time as "when the time out we change the roles". There is no other mention of navigational resources in the audio recording. This is interpreted as learners timing themselves but not necessarily with the screen-based timer or 'foregrounding' the screen-based timer and 'backgrounding' the navigational icons.

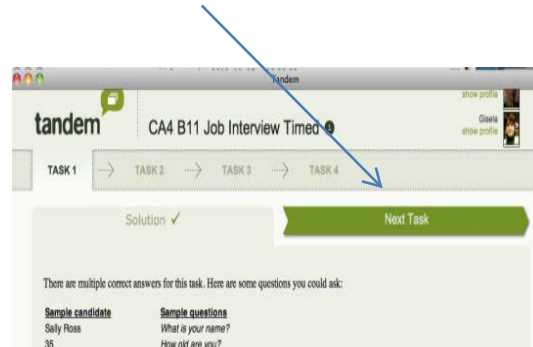
Cases 1 and 3 verbally mentioned the pop-up ‘Time Up!’/‘close’ button and the ‘next task’ resource as shown in figure 1 and figure 2, although not both times they (would have) appeared on the screen. This observation leads to mediation with navigational resources.

Navigational resources as invites for acceptance

Fig. 1 The pop-up ‘Time Up!’/‘close’ button (dynamic)



Fig. 2 Next Task button (static)



Both the pop-up ‘Time Up!’/‘close’ button and Next Task button (static) require navigation on the part of individual learners alongside the need for oral negotiation as to when learners navigate together. We consider the ‘close’ button on the Time-up! pop-up and the ‘next task’ button to be ‘invites’ which learners have to ‘accept’ by clicking them physically. The completed action is performed individually and multimodally: visually/textually from the computer (with the block of colour and then text; ‘close’ and ‘Next Task’) and somatically by individuals physically clicking these. The somatic mode relates to the body e.g. touch. Both cases 1 and 3 accompany this process with speech. This suggests that an initiation/response sequence may be possible to construct between learners and screen-based resources as ‘agents’ (Dourish and Belotti, 1993; Clark and Brennan, 1991).

The pop-up ‘Time Up!’/‘close’ button in particular, which would have occurred twice during the two tasks, would have surprised learners, yet no explicit reference or utterance of surprise was made by cases 2, 4, 5 and 6. From this we deduced that their trajectory with the navigational resources was somewhat different to cases 1 and 3. A possible explanation was that because the participants were already familiar with the tool and its pop-up features, their use had become normalised in their practice. Alternatively, they had already familiarised themselves with the resources prior to starting the task. We propose that the latter was the case as the tool is not used as common practice throughout the course so the students are not familiar with it.

The finding that the trajectories of 2, 4, 5 and 6 were different to that of cases 1 and 3 was triangulated through other sources and methods including the Tandem logs, a focus of time spent on task and the number of seconds before starting a new oral turn with their partner. The Tandem logs, collected from approximately 50 percent of the cases, indicated that the learners not only used navigational resources to respond to screen-based invites but to exit and re-enter the tool. We conceptualise the learners’ response to these navigational ‘requests’ as lower level actions which, when combined with other

completed actions, lead to the higher level action of managing the Tandem tool. The click of the button ('close' or 'next task') can be understood as 'accept' going forward as a completed (computer-human) turn for cases 1 and 3. However, we propose that the other cases by entering and exiting the tool, navigating back and forth, overrides the initial meaning of 'accept', potentially reassigning a different meaning to this resource.

The amount of time that learners spend completing the two tasks confirms the different trajectories. Cases 1 and 3 spent the minimum amount of time required for task completion according to task design with the pop-up and timers (11 minutes), whereas cases 2, 4, 5 and 6 spent less than the required time as can be seen in Table 2.

Table 2. Time-on-tasks 1 and 2

Case	Time required to complete tasks 1 and 2	Time taken by cases
Case 1	11 minutes	14:33
Case 2		8:07
Case 3		11:48
Case 4		3:33
Case 5		6:20
Case 6		9:31

Cases 2, 4, 5 appear to have self-regulated their time-on-task which is arguably easier to do without the presence of a timer and 'interruption' of a pop-up. Case 6 however, do mention that the time is running out ("when the time out we change the roles") indicating orientation towards time (or the timer) but not the navigational screen-based resources including the pop-up. It cannot be confirmed if case 6 used the screen-based timer or their own way of timing themselves to regulate the task. The behaviour of orientating to one indicator but not the others can be explained by the learners becoming familiar with the navigational resources and pop-ups previously to carrying out the (recorded) task.

These results suggest that whilst some learners are prepared to carry out the mediated action of task completion using the tool and resources according to task design (i.e. spontaneously), others are not. Some dyads appear to prefer to mediate with screen-based resources or tool before carrying out the recording (whether for linguistic reasons or reasons pertaining to the tool). Other cases, by choosing non-use of the navigational resources during the audio recording of the task and alongside the non-attention they pay to the pop-up 'Time Up!/'close' button suggests that they prefer to control the task conditions (controlling their own time and navigational moves). This highlights Norris' (2006; 2016) notion of simultaneous awareness/attention. By not paying attention to certain resources during the process of oral task completion learners are intentionally *backgrounding* their importance as they carry out the task orally and simultaneously. Reasons for mediating with the tool before task recording are discussed in authors (2017).

The trajectories were also triangulated by analysing the number of seconds before the start of the first turn with their partner, presented in table 3. This was because learners would have needed to read a number of textual information before they could start the oral task. Balaman and Sert (2017a) highlight the importance of long silences in their study as pointing to ongoing orientations to task interface. Reading these textual resources either aloud or silently would have taken time (text about purpose of the task; text about time limit; text warning about efficient use of time and text as lead-in to the sample

information). Therefore, if learners were reading (as a response to the screen), evidence of long silences would be apparent in the recordings.

Table 3. Number of seconds before start of talk

Case 1	23 seconds
Case 2	1 second
Case 3	28 seconds
Case 4	1 second
Case 5	1 second
Case 6	3 seconds

Table 3. shows that case 1 and 3 have the longest amount of time before beginning verbal turn-taking. We deduce that in their silence they are orientating towards or *foregrounding* the screen-based resources and therefore orientating away from or *backgrounding* their peer, highlighting learners shifting attention. We propose that long silences are potential indications of orientations towards the screen (Balaman and Sert, 2017) and responding in some way, including turns with specific screen-based resources.

In contrast, cases 2, 4, 5 and 6 start spoken interaction almost straight away indicating that they had no need to attend to the textual instructions and therefore were *backgrounding* these resources. We propose this is because they had pre-read them on (a) previous entry/ies to the tool.

In addition to the navigational resources' invites and the human physical responses to them which were sometimes accompanied by speech, another initiation/response pattern was identified, namely the pedagogical task instructions as a request to act. Responses to this were labelled as 'etic' in Table 1.

### **Pedagogical task instructions and prompts**

In Task 1 all cases responded to two pedagogical task instructions namely: *Use it to create questions and find out some important information about Student B* (with the 'it' referring to the sample information) and *Ask your partner at least five questions* as shown in figure 3. We consider the pedagogical task instructions to be a screen-based textual/visual 'initiation' that requires a visual/oral 'response' from learners because it is a 'request' to act. Learners' responses were in the form of a series of oral turns in their respective roles.

Fig. 3 Textual/visual resources

**Textual instructions:**  
*'Use it to create questions...'*  
 and  
*'Ask your partner at least 5 questions'*

**Textual/visual prompts**

The screenshot shows the following content:

**tandem** CA4 B11 Job Interview Timed

TASK 1 → TASK 2 → TASK 3 → TASK 4

See Solution

TIEMPO RESTANTE DE LA TAREA 03:51

This activity is a role play where Student A will pretend to be a **job consultant** (the interviewer) and Student B will be a **job applicant** (the interviewee).  
 In this first task, Student A (the interviewer) needs to find out some information about Student B (the interviewee).  
 Time limit: 4 minutes.  
**ATTENTION: You and your partner have exactly 4 minutes to complete the task. Use your time efficiently!**

Here is some sample information from a previous job applicant.  
 Use it to create questions and find out some important information about Student B. (Two sample questions are provided.)

Sample candidate	Your partner (sample questions)
Sally Ross	???
35	???
Boston	Where do you live?
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible, Friendly,	What words describe you?
Driver's license	???
40,000€ a year	???
9 a...	???

All interviewers asked at least 5 questions- (which one of the instructions explicitly requested), except for case 5 as shown in Table 4. Notably, many interviewers asked nearer 10 questions rather than the initial request of 5. The reason for this may be that the textual instruction to ask at least 5 questions was followed underneath by 11 textual/visual prompts for the interviewer (sample candidate and sample questions) as shown in figure 3. Therefore, the number of learner turns as a response to the instruction may have been shaped more by the number of available textual/visual prompts on screen over the textual instruction/request to create 5 questions. Rather than having 'reinforcing roles' (Kress and van Leeuwen, 2001) in the interaction between modes, the 11 prompts may have overrided the textual instruction requesting with respect to the amount of questions that were asked.

Table 4. Number of Questions asked by interviewer in Task 1

Case	Number of Questions asked by interviewer
1	12
2	9
3	9
4	8
5	4
6	9

The interviewers use the sample information to create questions as oral turns (as requested) which can be seen in figure 4 where the interviewer is beginning to ask questions. Bold in the transcripts indicates words with lexical similarity to the screen-based text or as emerging from mediation with the ‘?’. ‘Others’ as text (Radauskoski, 1999) are used as a conversational resource for creating oral turns. We explore this relationship in more depth in another paper (authors, forthcoming).

Fig. 4 Textual prompts as a conversational resource in the creation of oral turns  
Case 3 (Task 1)

<b>Ask your partner at least 5 questions.</b>	
<b><u>Sample candidate</u></b>	<b><u>Your partner (sample questions)</u></b>
Sally Ross	???
35	???
Boston	<i>Where do you live?</i>
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible. Friendly.	<i>What words describe you?</i>
Driver's license	???
40,000€ a year	???
9 am - 5 pm	???
Current Call	???

Example excerpt case 3

- A Okay, **How old are you?**  
 L I'm a thirty-five, sorry thirty-seven years old.  
 A **Where do you live?**

Whereas in task 1 learners do not generally go beyond the scope of the sample information as a conversational resource, this differs in task 2. In Task 2, learners are requested to respond to the textual instruction relating to two jobs presented: *Briefly describe them to your partner. Then answer your partner's questions about the jobs (if you don't know, be creative)*. Different interviewers describe each job orally using the details of each job presented textually on the screen (to a greater or lesser extent) as exemplified in figure 5. Again, the words in bold on the transcripts correspond lexically to those on the screen.



Fig. 5 Text as a conversational resource in the creation of oral turns Case 2 (Task 2)

Here are two different jobs you have available. Briefly **describe both of them** to your partner. Then **answer your partner's questions about the jobs**. (If you don't know the answers, be creative!)

<b>Job 1: Bilingual Sales Manager</b>	<b>Job 2: Travel agent</b>
London, England, New York, NY, or Madrid, Spain	Any major city in Europe
Must be willing to travel to other cities once a month	Must be willing to travel to other cities once a month
High school diploma required	High school diploma and business school required
Spanish and English required	English required; a foreign language (Spanish, French or Chinese) is highly desired
10+ years of sales experience required	2 years experience required in the travel industry or in business
Microsoft Word, Excel, Powerpoint and Access required	Microsoft Word, Excel, and Powerpoint required
Must be a good salesperson	Must be a good communicator
9 a.m. - 5 p.m. or 12 - 8 p.m.	9 a.m. - 6 p.m.
Must have a car	No car required
€30,000 a year + commission	30,000€ a year

#### Excerpt

- G Well err there is enough for me. Err I can offer you two works in our company. Er, the first is a **bilingual sales manager** and did require long work experience in that type of sales. Also you said you work two years only. I thing you could be a good candidate. The (pause) second job (pause) is in like a **travel agency**, Um the job is a office in a **European city**, but in this moment I don't know where is the vacant of the travel agency. Um speaking **English is required** but it is also important to speak another **language** fluently like **Spanish or French**. The **experience** is a little few; **two years** is good and I don't know what more explain you.

In Task 2, Cases 1, 2, 3 and 6 describe the details relating to the two jobs presented (Bilingual Sales Manager and the Travel Agent) as an extended oral turn. However, the interviewer for case 4 presents her partner with two jobs that are not part of the task design: a primary school teacher and secondary school teacher of English and positions herself as a headmistress of the school who is calling by telephone. Similarly, case 5 presents their partner with two jobs: a receptionist night technician who does manicures and a personal trainer. This means that cases 4 and 5 have created their own jobs – possibly responding to the textual instruction in the task to ‘be creative’, although this cannot be confirmed. These cases completely *background* the jobs presented to them textually on the screen by not attending to the job descriptions provided. This result highlights that through mediation learners can decide on the importance of one textual instruction over another, which may be designed to complement each other in the initial task design but which learners can reconfigure with their own agency, depending on the importance or (not) they give to it. Whilst learners are responding to textual instructions in a series of oral turns they are also responding to the textual prompts (or not) as a conversational resource.

## 8. Results

The aim of the study was to understand how learners carried out mediated action as task completion through and/or with the available semiotic resources in an online spoken interaction task.

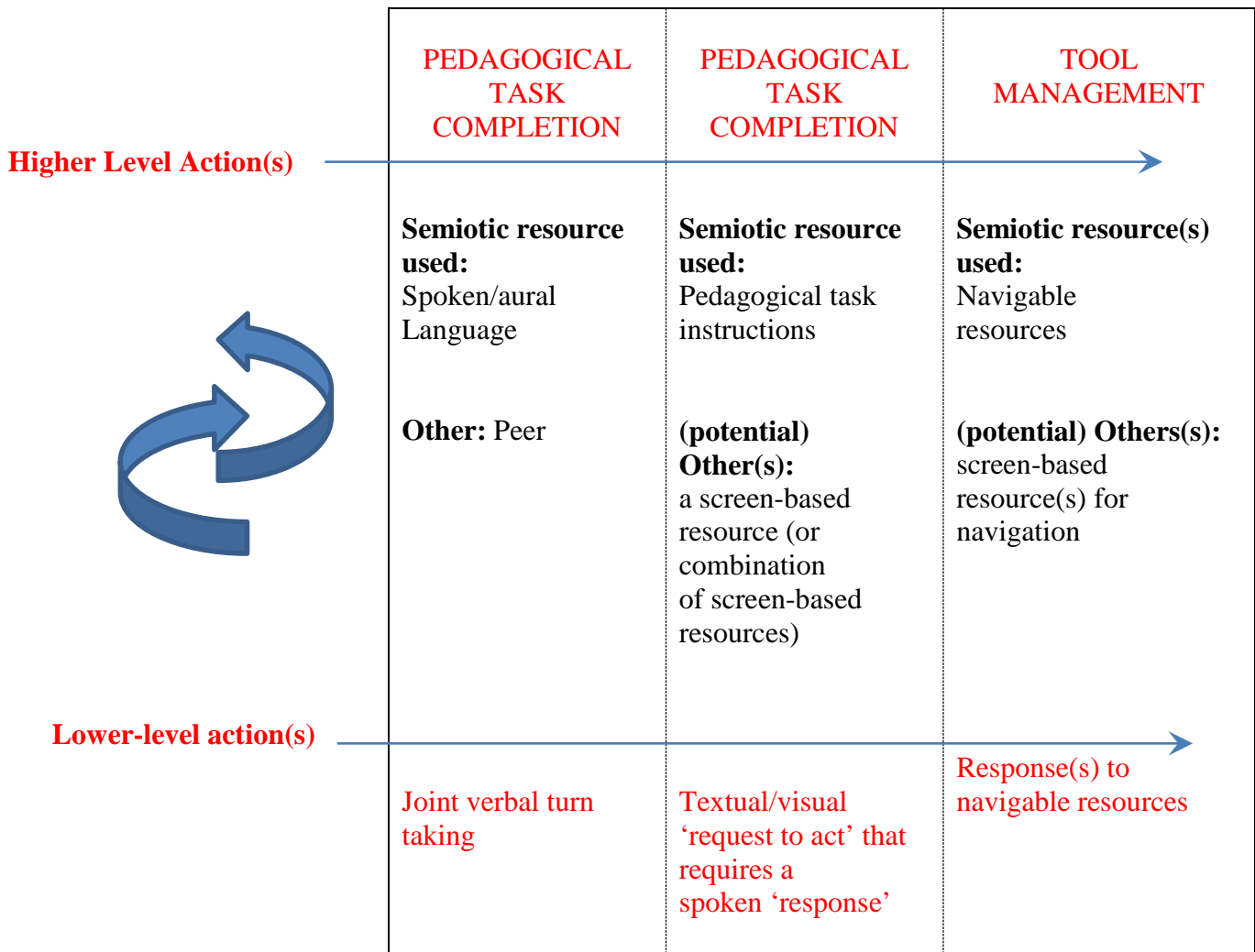
‘Semiotic initiation and response’ (Coffin and Donohue, 2014) has been a useful notion in characterising the mediated action for task completion. There is an overarching mediated action that is being carried out between task instructions (a textual screen-based ‘request’) and learners (as a series of oral turns as a ‘response’) whilst carrying out oral turns with their partner. In addition, learners’ encounter with the pedagogical task instructions and navigational resources shows how the sense making is constructed one-sidedly, rather than co-constructed, making the learner solely responsible for the emerging meaning, following Raudaskoski (1999).

Based on the results, we present the following framework that summarises the findings. The framework draws on Norris’ (2004; 2006) notion of *higher* and *lower level actions* to highlight how actions in the mediation process may have two distinct but entwined goals in which the learners’ actions appear to be directed towards. The *higher level actions*, as large scale actions that learners achieve in the task process, are the completion of the pedagogical task and the management of the tool. These two *higher level actions* are made up of a multiplicity of chained *lower level actions* which are smaller-scale actions.

The pedagogical task completion is achieved through the *lower level actions* of verbal turns, jointly co-constructed between peers as well as what can be understood as the response (in the form of a series of turns) to pedagogical task instructions (e.g. create 5 questions) that request them to act.

The *higher level action* of (technological) tool management is achieved through learners’ use of the navigational resources as a chain of smaller-scale lower level actions which are both individually (physically) and jointly accomplished and sometimes orally negotiated. The *lower level action* of navigation means that learners can reconfigure (navigate back and out of the tool and re-entering- not just navigating forward) and repurpose (retrieval of task information and instructions without being navigated or pressured by potential pop-ups and timers) the original task and tool design.

Figure 6. Higher and Lower Level Actions in an audioconferencing task



The framework highlights that learners are ‘multi-tasking’ during tasks (Knight and Barbera, 2016) as learners’ carry out oral turns and initiate and respond to screen-based resources.

## 8. Conclusions

The results reveal that audio-conferencing as a mode-as-channel of communication is not necessarily “voice, but no image” (Yamada, 2009) but that learners can be faced with a variety of visual/textual/navigational screen-based resources as a ‘multiple ensemble’ (Bezemer and Kress, 2016) whereby they use the screen-based resources to carry out the mediated action. Whilst semiotic mediation for pedagogical task completion appears relatively straight forward (spoken turns as lower level actions are used to achieve higher level action of pedagogical task completion), the mediational process is more complex for tool management. It is not only a multimodal process involving different oral turn-taking combinations but a multisensory process where learners attend to the screen through their sensory (visual) system and motor system as they touch, responding to screen-based resources that also initiate action. Learner agency is manifest not only

through how learners carry out the task through oral, visual and somatic initiations or responses but also through the importance and non-importance given to the textual instructions and prompts as well as through learners' repurposing of the navigational resources.

The concepts of *lower* and *higher level actions* (Norris, 2004; 2006) have allowed for an analysis of 'tools' as semiotic resources for mediation *within* a technological tool without losing sight of their interconnection. This is important for two reasons: it adds to the focus on tool affordances regarding the purpose(s) of a tool and task design by supporting the analysis of mediational processes that may or may not form a part of that design purpose. Furthermore, the notions of *lower* and *higher level actions* support an understanding of how learner agency may occur at lower levels (in the process of mediating with screen-based resources and using the verbal turn-taking system) whilst simultaneously supporting higher level actions of tool management and pedagogical task completion.

In addition, what appears to be the avoidance of certain types of experience by some students whilst speaking (e.g. being navigated, being pressured by time) suggests that there may be some experiences that some learners prefer to avoid. For online learners in particular, two of the affordances of online learning are that learners can control their own time and that they have navigational freedom. Therefore, tools which intend to control both 'affordances' may be a specific problematic area for tool and/or task design. In this light, it is important not to assume that multimodal experiences will be positive and rich for learners but keep in mind the 'dual' goal of techno-pedagogical task completion together with the perceived acceptable conditions online learners might have when they complete these goals.

Finally, terms that suggest that CMC is being carried out through monomodal channels such as videoconferencing/audioconferencing/text-chat, on closer inspection reveal this is increasingly less so (e.g. Liddicoat, 2010; Lamy, 2006) and therefore potentially masks over other screen-based modes learners may be faced with.

## **9. Future research and implications**

Further research is needed that takes into account the growing importance of the screen-based resources in language learning. This study confirms that learners are attending to both their peer and the screen. A purely lingual lens is therefore not sufficient to capture how learners may be positioned as semiotic 'initiators' and 'responders' (Coffin and Donohue, 2014) in this process.

Conceptualising screen-based resources as 'others' more relevant in their support for turn-taking through non-verbal means (i.e. visually/textually/somatically), by attending to the screen interface's materiality is therefore important in order to see how and if learners respond to them as turns or not. Whilst the emic perspective allowed an identification of screen-based resources as conversational resources, the etic perspective highlighted the presence of navigational resources as 'agents' that could initiate or respond. Therefore, future research might attend not only to "the silent visible displays of the hearer work" (Goodwin, 2013 p. 8) of participants (in oral interaction) in order to understand the co-operative social organization involved in a single shared action but also to the silent visible displays of the screen. Mediation with 'others' (Radauskoski, 1999)

as signs and text/discourse as conversational resources (e.g. Kraut, Fussell and Siegel, 2003) can be considered being part of a semiotic field or 'layer' of action (Goodwin, 2013). However, the study shows how resources can also act as or be responded to as a conversational agent. How these resources are orientated to and how they shape talk in this 'encounter' (Radauskoski, 1999) needs further exploration.

Whilst combining both emic/etic perspectives has provided fuller insight into both human-human and human-computer interaction. However, this is not without limitations. One major limitation is that we cannot see whether learners' orientations are happening in real time, verbal reports of their previous actions or whether these orientations co-occur with talk minute-by-minute. However, because the aim of the study is to establish whether mediation with screen-based resources occur in the first place-and if so, which ones, rather than the minute-by-minute account, we take the discourse as a 'fingerprint' alongside the triangulation of sources to be robust enough to answer this particular research question. Whilst it is true that eye-tracking technologies combined with micro analytic discourse offers an arguably more accurate understanding of the minute-by-minute action, online learners who are not easily reached by such methods or prefer that their pre-recording behaviours remain hidden may need approaches such as the one we use. The incorporation of learner questionnaires about pre-recording behaviours and think-aloud protocols, would increase validity.

There are a number of implications for task design. There is the consideration as to whether screen-based resources *represented* to learners have the purposes to prompt or shape spoken turns or as a means for the teacher/designer to *communicate* to students (which may draw the learner's attention away from their peer). The 'interplay of modes' (Schnotz, 1999) is also important, particularly as one result tentatively suggests that the textual/visual prompts may be able to support an increase in number of oral turns taken than suggested in the instructions- a common goal for spoken interaction tasks.

## Bibliography

- Balaman, U., & Sert, O. (2017a). Local Contingencies in L2 Tasks: A comparison of Context-Sensitive Interactional Achievements across Two different Task Types. *Bellaterra Journal of Teaching & Learning Language and Literature*, Vol.10(3), 9-27.
- Balaman, U., & Sert, O. (2017b). The coordination of online L2 interaction and orientations to task interface for epistemic progression. *Journal of Pragmatics*, 115, 115-129.
- Benson, P. (2015). You Tube as text: spoken interaction analysis and digital discourse. In Jones, R. H., Chik, A., & Hafner, C. A. (Eds.), *Discourse and digital practices: doing discourse analysis in the digital age* (pp. 81- 96). Routledge: Oxford and New York.
- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: a social semiotic frame*. London: Routledge.
- Blin, F., & Jalkanen, J. (2014). Designing for language learning: agency and languaging in hybrid environments. *APPLES : journal of applied language studies*, 8 (1), 147-170. Retrieved from <http://apples.jyu.fi/ArticleFile/download/433>
- Canto, S., de Graaff, R., & Jauregi, K. (2014). Collaborative tasks for negotiation of intercultural meaning in virtual worlds and video-web communication. In M. Gonzalez-Lloret, & L. Ortega (Eds.), *Technology and tasks: Exploring technology-mediated TBLT* (pp. 183-212). Washington, DC: Georgetown University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition*, 13(1991), 127-149.
- Coffin, C., & Donohue, J. (2014). A language as social semiotic-based approach to teaching and learning in Higher Education, *Language Learning*, 64(1), 1-212.
- Collentine, K. (2009). Learner use of holistic language units in multimodal, task-based synchronous computer-mediated communication. *Language Learning & Technology*, 13(2), 68-87.
- Dourish, P., Bellotti, V., Mackay, W., & Ma, C. Y. (1993, December). Information and context: lessons from the study of two shared information systems. In *Proceedings of the conference on Organizational computing systems* (pp. 42-51). ACM.
- Flewitt R (2008) Multimodal literacies. In Marsh, J. and Hallet, E. (eds), *Desirable Literacies: Approaches to Language and Literacy in the Early Years*. London: Sage, 122–139.
- Flewitt, R., Hampel, R., Hauck, M., Lancaster and Jewitt, C. (2016). Multimodal data and transcription. In Jewitt, C. (Ed), *The Routledge Handbook of multimodal analysis*. London: Routledge. 40-53.

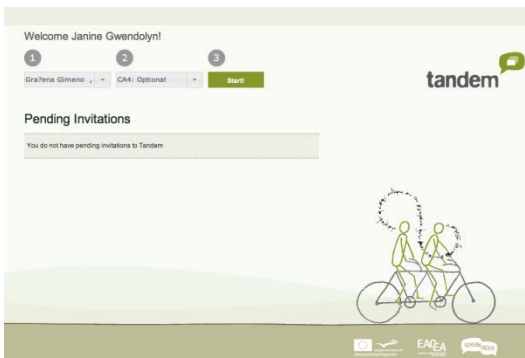
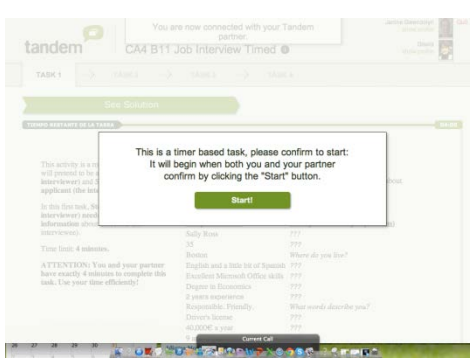
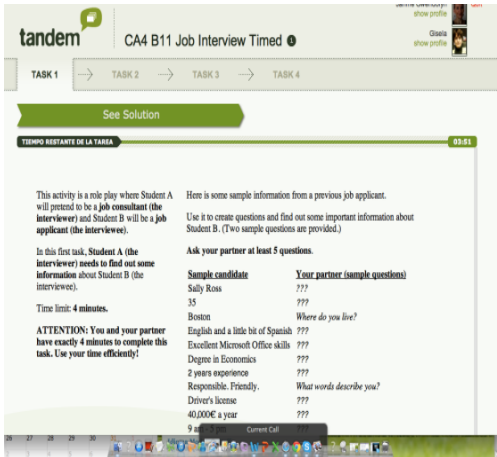
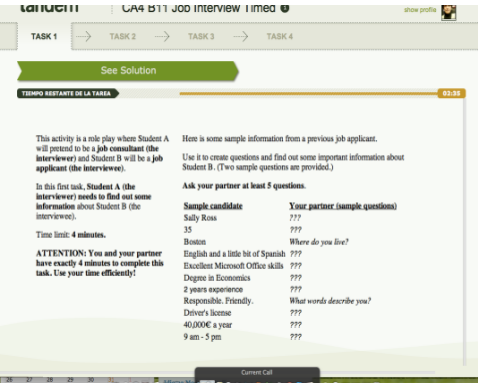
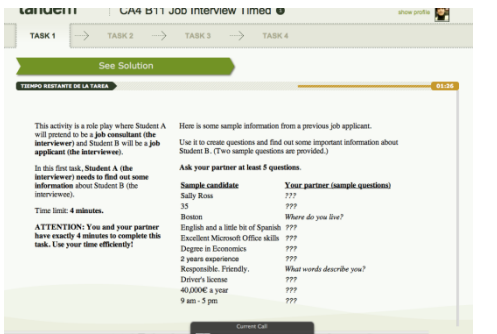
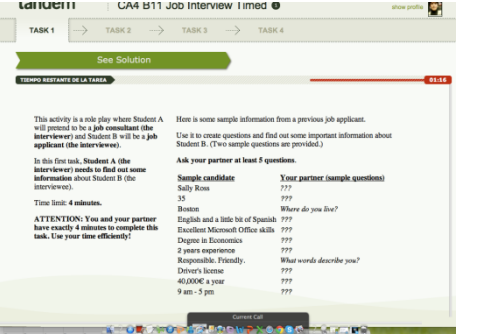
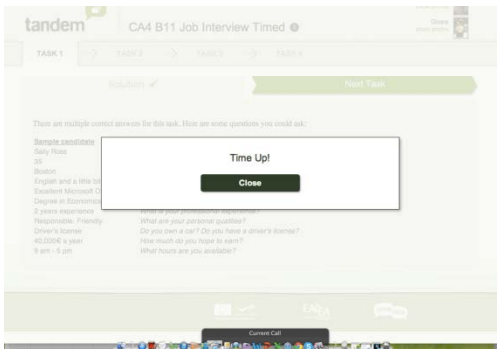
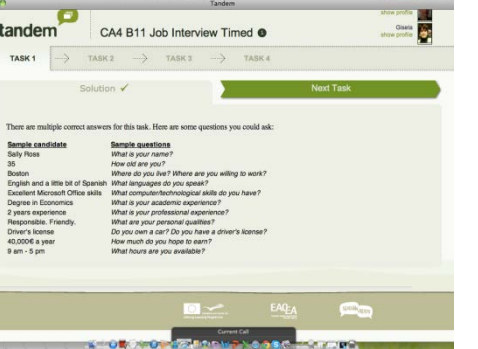
- Gilabert, R., Manchón, R., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117-135.
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46(1), 8-23.
- Hampel, R., & Hauck, M. (2006). Computer-mediated language learning: Making meaning in multimodal virtual learning spaces. *The JALT CALL Journal*, 2(2), 3-18.
- Hampel, R. (2010). Task design for a virtual learning environment in a distant language course. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 131–153). London: Continuum.
- Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(02), 116-137.
- Hauck, M. (2010). The enactment of task design in telecollaboration 2.0. In Thomas, M., & Reinders, H. (Eds.), *Task-based language learning and teaching with technology* (pp. 197-217). London: Continuum International Publishing group.
- Helm, F., & Dooly, M. (2017). Challenges in Transcribing Multimodal Data: A Case Study. *Language Learning & Technology*, 21(1), 166-185.
- Herring, S. (2015). New Frontiers in Interactive Multimodal Communication. In Alexandra Georgakopoulou and Tereza Spilioti (eds.). *The Routledge Handbook of Language and Digital Communication* (pp. 398- 402). New York: Routledge.
- Jenks, C. (2014). *Social interaction in second language chat rooms*. Edinburgh University Press.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing Multimodality*. London: Routledge
- Knight, J., & Barbera, E. (2016). The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks. In Pixel (Ed.) *Proceedings of the 9th International Conference on ICT for Language Learning* (pp. 248-252). Florence, Italy, 17-18 November, 2016. Padova, Italy: Libreriauniversitaria.it
- authors. (2017). Learner agency and its effect on spoken interaction time in the target language. *JALT CALL Journal*, 13(3).
- authors, (2018). But the computer say me the time is up: the shaping of oral turns mediated with and through the screen. Manuscript submitted for publication.
- Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-computer interaction*, 18(1), 13-49.

- Kress, G. (2003). *Literacy in the new media age*. London: Routledge
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Edward Arnold.
- Lantolf, J. P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford, England: Oxford University Press.
- Lamy, M.-N. (2006). Multimodality in second language conversations online: Looking for a methodology. In proceedings of third international conference on multimodality (pp. 385- 403). Pavia, Italy.
- Liddicoat, A. J. (2010). Enacting participation: Hybrid modalities in on-line video conversation. In C. Develotte, R. Kern, & M. N. Lamy (Eds.), *Décrire la conversation en ligne* (pp. 37–50). Lyon, France: ENS Editions.
- Liou, H. (2011). Blogging, collaborative writing, and multimodal literacy in an EFL context. *WorldCALL International Perspectives on Computer-Assisted Language Learning*, 3-18.
- Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse studies*, 9(2), 194-225.
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. London: Routledge.
- Norris, S. (2006). Multiparty interaction: a multimodal perspective on relevance. *Discourse studies*, 8(3), 401-421.
- Norris, S. (2016). Concepts in multimodal discourse analysis with examples from video conferencing. In *Yearbook of the Poznan Linguistic Meeting* (Vol. 2, No. 1, pp. 141-165). De Gruyter Open.
- Oben, B., & Brône, G. (2015). What you see is what you do: on the relationship between gaze and gesture in multimodal alignment. *Language and Cognition*, 7(4), 546-562.
- O'Halloran, K. L. & Smith, B. A. (ed) (2011). *Multimodal Studies: Exploring Issues and Domains*. New York & London: Routledge.
- Raudaskoski, P. (1999). *The use of communicative resources in language technology environments. A conversation analytic approach to semiosis at computer media* Doctoral dissertation. Retrieved from <http://vbn.aau.dk/ws/files/72280035/pirkkosphd.pdf>
- Rossolatos, G. (2015). Taking the “multimodal turn” in interpreting consumption experiences. *Consumption Markets & Culture*, 18(5), 427-446.
- Sacks, H., Schegloff, E.A. and Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation, *Language*, 50(4):696-735.



- Schnotz, W. (1999). Visual learning with new technologies: Introduction. *European Journal of Psychology of Education, 14*, 163-165.
- Tudini, V. (2014). Conversation Analysis of computer-mediated interactions. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*, 1–7. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. Lantolf (Ed.), *Sociocultural theory and second language acquisition* (pp. 245-259). Oxford, England: Oxford University Press.
- Vetter, A., & Chanier, T. (2006). Supporting oral production for professional purposes in synchronous communication with heterogenous learners. *ReCALL, 18*(01), 5-23.
- Vygotsky, L. S. (1981). The instrumental method in psychology. In J. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 134- 143). Armonk, NY: Sharpe.
- Wertsch J. V. (1998) *Mind as Action*, Oxford University Press, Oxford.
- Yamada, M. (2009). The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: Experimental study. *Computers & Education, 52*(4), 820-833.

# Appendix A: Example screenshots of reconstructed-task simulation by researcher

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8. 

## Appendix B: Labelling of The multimodal ensemble of interface pages (Task 1)

**Visual/textual resources:**  
Information about task sequence

**Visual/textual/navigational**  
'See solution'

**Visual resources (dynamic)**  
Timer, pop-up

**Textual/visual resources:**  
Information about roles  
Information about the communicative purpose  
Instructions  
Prompts  
Information: time limit with warning

The screenshot shows the Tandem interface for 'CA4 B11 Job Interview Timed'. At the top, there is a task sequence: TASK 1 → TASK 2 → TASK 3 → TASK 4. Below this is a green bar with 'See Solution' and a timer showing 'TIEMPO RESTANTE DE LA TAREA' at 03:51. The main content area is divided into two columns. The left column contains instructions for the role-play activity, including the roles of Student A (job consultant) and Student B (job applicant), and a 4-minute time limit. The right column contains sample information from a previous job applicant, including name, age, location, language skills, education, experience, personality, and income. Below this is a table with 'Sample candidate' and 'Your partner (sample questions)'. The table lists Sally Ross, 35, Boston, English and a little bit of Spanish, Excellent Microsoft Office skills, Degree in Economics, 2 years experience, Responsible, Friendly, Driver's license, and 40,000€ a year. The sample questions include 'Where do you live?', 'What words describe you?', and 'What are your personal qualities?'. A 'Close' button is visible at the bottom of the content area.

## Interface page (when time is up for Task 1)

**Visual/textual/navigable resources**  
Time Up! Pop-up (visual/textual)/close (navigational)

**Visual/textual/navigational**  
'Next task' (phased out) and 'start' button

The screenshot shows the Tandem interface when time is up for Task 1. A 'Time Up!' pop-up window is displayed over the task content. The pop-up window has a 'Close' button. The 'Next Task' button is visible and appears to be phased out. The background content is dimmed. A 'Close' button is visible at the bottom of the content area.

## Appendix C: Analysis of screen-based resources

No screenshot	Waiting for confirmation
<b>Textual/visual resources (static):</b> screenshots 3-6	<p>INTERFACE PAGE 1 ON SCREEN</p> <ul style="list-style-type: none"> <li>-‘Task 1’ foregrounded (Task, 2, 3 and 4 backgrounded)</li> <li>- about roles or Student A or B</li> <li>- about purpose of the task</li> <li>-about time limit</li> <li>-warning about efficient use of time</li> <li>-lead-in to sample information</li> <li>- create own questions + symbol ?</li> <li>-instruction to ask the minimum of 5 questions</li> <li>- Text from sample candidate</li> <li>- Textual sample questions e.g ‘Where do you live?’</li> </ul>
<b>Visual (dynamic):</b> screenshots 3-6	-The visual timer counting the seconds left and changes colour with ‘ <i>Tiempo Restante de la Tarea</i> ’ (Time left for task) banner
<b>Visual/textual/navigable:</b> screenshots 3-6	-‘See Solution’ banner
<b>Visual/textual/navigable (dynamic)</b> screenshots 7 screenshot 8	<p>SCREEN (INTERFACE PAGE 1 BACKGROUNDED)</p> <ul style="list-style-type: none"> <li>-A pop-up sign indicating ‘Time up!’ and ‘Close’</li> </ul> <p>INTERFACE PAGE 2 ‘SOLUTIONS PAGE’</p>
<b>Textual (static)</b> screenshot 8 screenshot 8	<ul style="list-style-type: none"> <li>-information that there are multiple correct answers</li> <li>-information of sample candidate and sample questions</li> </ul>
<b>Visual/textual/ Navigable</b> screenshot 8	-‘Next task’ banner
screenshot 9	SCREEN (INTERFACE PAGE 2 BACKGROUNDED)
<b>Visual/textual/navigable</b> screenshot 9	<ul style="list-style-type: none"> <li>-information that it is a timed task request to start by clicking the ‘Start’ button</li> </ul>
screenshot 10-12	INTERFACE PAGE 3
<b>Textual/visual</b> screenshot 10-12	<ul style="list-style-type: none"> <li>-‘Task 2’ foregrounded (Task 1, 3 and 4 backgrounded)</li> <li>-information that two jobs are available and the need to explain them</li> <li>-information that partner should ask 3 questions and say which they prefer</li> <li>-information about time limit</li> <li>-instruction to describe the two jobs to partner, answer partner’s questions and be creative if you don’t know the answers</li> <li>Text describing two jobs</li> </ul>
<b>Visual (dynamic):</b> screenshot 10-12	-The visual timer counting the seconds left and changes colour with ‘ <i>Tiempo Restante de la Tarea</i> ’ banner
<b>visual/textual/navigable (dynamic)</b> Screenshot 13	<p>SCREEN (INTERFACE PAGE 2 BACKGROUNDED)</p> <ul style="list-style-type: none"> <li>-A pop-up sign indicating ‘Time up!’ and ‘Close’</li> </ul>
<b>Visual/textual/navigable</b>	-‘Next Task’ Banner



17.4

Paper 5: Multimodal meaning making: navigational acts in an online speaking task

## **Multimodal meaning making: navigational acts in an online speaking task**

### **ABSTRACT**

Intentionally clicking screen-based navigational resources can be one way in which learners exercise agency (authors, 2017). Because such navigational acts require learners to be 'semiotic initiators and responders' (Coffin and Donohue, 2014) of navigational resources during tasks, possibilities may exist for meaning making beyond the purely verbal. However, their meaning making potential and impact on task processes have received little attention in the field of Second Language Acquisition.

This case study explores how dyads from two different data sets, across three online speaking tasks (information gap, opinion sharing, role-play), carry out navigational acts facilitated through audio-conferencing. Learners' explicit mention of resources on the screen was used to 'track' their trajectories during task process. Multimodal (inter)actional analysis (Norris, 2004) and computer mediated discourse analysis (Herring, 2004) was used for the analysis with results triangulated using different data sources and methods.

Results suggest that depending on different case trajectories, learners can orally negotiate navigational acts contributing to peer-to-peer meta-modal talk, or mediate with them in such a way that this oral negotiation is eliminated from talk in the target language. Furthermore, the non-verbal meaning making aspects that were negotiated related to technological tool-use and learner roles as tool user or tool manager.

**Key words:** navigation, learner agency, multimodal, discourse, non-verbal meaning making, negotiation, spoken interaction, screen-based resources, task

## 1. Introduction

Carrying out navigational acts, understood as learners' intentional actions for navigational purposes, such as clicking a screen-based navigational resource, can form part of how language learners exercise their agency in online tasks. Navigation can be defined as "the process of determining and maintaining a course or trajectory from one place to another" (Gallistel, 1990, p.35). The ability to maintain a trajectory is a core component of the process involved in approaching a destination (Ferretti, Adornetti, Consentino and Marini, 2013) and goal achievement. Therefore, navigation implies learners making choices about how to achieve particular goals and learners' actions based on those choices are needed in order to maintain the trajectory for goal achievement. To maintain a trajectory, learners also need to control navigational possibilities. Learners' choices and intentional actions not only can shape task outcomes (Knight, Barberà and Appel, 2017) but also conceivably offer ways for learners to shape the task process as well as the non-verbal meaning that may emerge from this scenario.

Parallel to this scenario is the fact that digital tools that facilitate computer-mediated communication (henceforth CMC) are expanding in range and complexity. For example, through tools such as Skype, Google Hangout and Webex, learners may communicate through audio, video and/or text whilst also carrying out navigational acts (e.g. click on a hyperlink or screen-based navigational resource that moves the task on in some way). This scenario



conceivably reflects an “ever expanding semiotic budget” (Blin and Jalkanen, 2014, p.150) whereby learners can be ‘semiotic initiators and responders’ (Coffin and Donohue, 2014) not only with language as a semiotic resource but also with navigational screen-based resources (e.g. button for accepting an invite from another participant to connect).

Navigational screen-based resources, when considered as forming part of all the semiotic resources learners have at their disposal during a task, may have what van Leeuwen calls ‘meaning potential’ (2004, p.3). However, their meaning potential realised through learners’ choices and related actions and their effect on spoken interaction, have received little attention in Second Language Acquisition research (henceforth SLA). Whilst a dominant conceptualisation of meaning making during task processes has been studied through the notion of ‘negotiation for meaning’ (henceforth NfM) some argue that the term is problematic because of its dominance and narrow focus that has led to an omission of other understandings (Block, 2003). Furthermore, the term ‘Negotiation of Meaning’ has been applied to purely verbal meaning making over non-verbal, highlighting ‘the lingual bias’ (Block, 2013) in the field. However, the presence of navigational screen-based resources in tasks and the acts that learners can carry out using them can be another way in which learners might negotiate meaning. For example, a navigational resource may be designed to move learners forward but learners may use it to go backwards and therefore resignify its original purpose/meaning.

In this study, we seek to explore non-verbal meaning negotiation that pertains to navigational acts. In order to advance our understanding of non-verbal meaning making and to gain insight into how navigational acts may also be shaping spoken interaction in online language learning environments, we use a multimodal perspective. This perspective has the aim of encompassing some of the complexity of all the layers and modes involved in meaning creation (Calvo Ferrer, Melchor-Couto and Jauregui, 2016). It signals a shift away from a purely linguistic approach to meaning making. Greater insight into screen-based navigational resources as having meaning making potential and their relationship with online spoken interaction may inform future task and/or tool design for task-based synchronous CMC whereby navigational screen-based resources play a prominent role in learners' experience(s).

Throughout the paper we use the term 'navigational resources' in place of 'screen-based navigational resources' for the sake of expediency.

## **2. Theoretical Framework**

### **2.1 Learner agency, Computer-Assisted Language Learning and tasks**

Without learner agency, navigation cannot take place as part of the mediation process. Agency is understood in this study as "the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives" (Martin, 2004, p. 135) and is used over the more common "the

socioculturally mediated capacity to act” (Ahearn, 2001, p.112). However, many agree that an individual’s capacity to act is socioculturally, contextually and interpersonally mediated (Mercer, 2011).

Whilst a few studies have focused on agency and tasks involving speech (Van Lier, 2008; Novick and Sutton, 1997), more recently Knight, Barberà and Appel (2017) identified types of agency in online speaking tasks. One type, related to navigational choices and acts, is named ‘directional agency’. Directional agency is described as being physical in nature, but also implies a spoken or written aspect accompanying or preceding it, necessary for decision-making. It involves choices and physical moves to navigate (e.g. with a button or hyperlink) using technological features (Knight, Barberà and Appel, 2017). Learners’ use of directional agency was found to affect both time interacting in the target language (henceforth TL) and whether spoken interaction was recursive or not (Knight, Barberà and Appel, 2017).

## **2.2 Navigation and CALL**

Navigation has been studied in CALL in relation to navigating hypermedia (visual/ musical/ animation elements) and/or hypertext (textual) in Intelligent CALL systems and tools. Heift, (2002) identified three different personas within an ICALL system, namely ‘browsers’, ‘peekers’, and ‘adamants’, which were reflected in navigational patterns. Results suggested that language proficiency and the amount of control the learner could exert in an instructional situation needed to be considered. Navigation as a design consideration has also been a

focus for ICALL (e.g. Virvou and Tsiriga, 2001; Amaral and Meurers, 2011). Virvou and Tsiriga's (2001) study on adaptive navigation support found that due to extra navigational freedom they provide, hyper-documents imposed greater cognitive loads on users compared to linear documents.

The integration of navigational resources in tasks is one way in which learners can be 'scaffolded', namely the steps taken to reduce the degrees of freedom in carrying out some tasks so that learners can concentrate on the difficult skill they have in the process of acquiring (Bruner, 1978). Whilst, 'scaffolding' in SLA has been considered with respect to language learning (e.g. Gibbons, 2002) the scaffolding of tool-use is also a research area. Studies in virtual environments such as Second Life have also discussed navigation (including proxemic acts) but these are not explored because of the scope and limitations of this paper.

In CMC tasks, navigation can also occur using hypermedia synchronously whilst communicating. Hypermedia can support individual navigation within and across sites and be used in a shared way (Riva, 2001). Shared hypermedia or SHY tools allow learners to communicate whilst synchronously browsing. In the field of CMC very few studies have focused on navigation and the relationship with speaking, possibly because many tools for spoken interaction have not all had hypermedia possibilities or navigational possibilities are not incorporated in task design. However, Blake (2016) notes that with CMC exchanges the computer and the screen mediate the entire communication experience and highlights studies that have warned that the interface profoundly affects conversational dynamics. Because navigation can be carried out through hypermedia and hypertext on interface pages it can be a task and tool design consideration.

In CMC, speaking is one of the most important skills that engages language teachers and students (Mahdi, 2014). Studies on audioconferencing and videoconferencing for speaking skills (see Blake, 2016 for overview) have not primarily focused on navigational aspects.

However, Levy and Kennedy's (2004) study highlighted the affordances of an audio-conferencing tool (with real-time audio and video, text-based chat, document sharing and whiteboard graphics exchange) that facilitated the ability for two learners to browse through web sites and see the same on-screen material as they talked about it. Appel, Robbins, Moré and Mullen (2012) focused on how navigational features of a task might guide the conversation and act as a 'scaffold' influencing the use of the TL for task completion. Results of learners using three versions of the interface (HTML version, Tandem with buttons and Tandem without confirmation buttons) revealed that learners reported feeling more nervous with the Tandem, while the HTML version allowed learners to prepare beforehand, leading to less naturalness and spontaneous talk.

In a telecollaboration learning project to develop speaking amongst other skills, Guth and Helm (2011) outlined three tasks using Skype for speaking (and textchat) alongside other web-based sources such as images, video, dictionaries and translation services to support discussion. In some tasks, learners switched from the Skype window to an internet browser or kept other tools open at the same time e.g. wikis (implying navigation). In the first session, learners spent a

lot of time becoming familiar with the technology with some students reportedly feeling anxious about it.

### **2.3 Semiotic Mediation, navigational screen-based resources and spoken interaction**

The notion of 'semiotic resources' or 'tools' (Vygotsky, 1981) is a way of understanding how learning is mediated through the use of printed materials, physical environment, gestures and classroom discourse (Lantolf, 2000). In online tasks, navigational resources can form a part of an "ever expanding semiotic budget" (Blin and Jalkanen, 2014, p.140) but they can also be a means whereby learners control their screen-based resources e.g. navigating between interface pages or minimising browser windows.

A number of studies on CMC for spoken interaction have highlighted how clicking various screen-based resources can form part of the accomplished interaction. Lamy (2012) found that in audiographic conferencing, a user may close a conversation by typing 'bye for now', by clicking a specific button, or withdrawing orally (Lamy, 2012) and proposes that clicking a resource that means 'temporarily away' is likely to influence the direction of the conversation differently compared to disconnecting altogether. Liddicoat (2010) found that in an opening videoconferencing exchange the securement of a non-present co-participant for the conversation was achieved by the initiator through the technology: a message via the computer (screen). According to Liddicoat (2010), this sequence resembled summons-answers sequences. Furthermore, whilst learners were

establishing a channel, learners orientated to the technological interface in their talk, rather than to personal interaction with their peer before moving to the first topic. In a synchronous CMC scenario both Cunningham, Fägersten and Holmsten (2010) and Fägersten, Holmsten and Cunningham (2010) identified meta-modal discourse pertaining to the screen-based resources of a seminar tool involving video, whiteboard facility and textchat. The meta-modal communication occurred in the form of students trying to help other participants with the technology (the pointer function, text box) (Cunningham, Fägersten and Holmsten, 2010). Fägersten, Holmsten and Cunningham (2010) found that the metamodal discourse helped interactants navigate the medium. Learners' overt attendance to (screen-based) communication modes was considered part of video-conferencing, serving to initiate repairwork and facilitate the progression of communication. Authors' (2016) study of peer-to-peer spoken interaction tasks using an audioconferencing tool found that learners were multi-tasking as they interacted with different screen-based resources. Oral negotiation of meaning was not limited to the pedagogical task. Focus included co-ordinating navigational resources, which learners acted upon and which shaped how and what was negotiated. However, the extent to which this occurred depended on if and how dyads followed the route/sequence of the task design. Some dyads had created their own route or 'lines of desire': an architectural term that can be applied to technology that highlight the paths people make, which are often shortcuts that ignore the given route (Luckin and Du Boulay, 2003).

## **2.4 Negotiating Meaning, online speaking tasks and multimodality**

In SLA, semiotic mediation from a sociocultural perspective has predominantly focused on the use of spoken or written language to mediate meaning making and language learning. However, the term 'negotiation for meaning' (henceforth NfM) from a cognitivist perspective has stemmed from research that seeks to determine which task variables have a positive effect on quantity of meaning negotiation through turn-taking (Foster and Ohta, 2005), with a focus on communication breakdowns (from Varonis and Gass, 1985). Some studies have used the term 'negotiation of meaning' but maintain the same construct as NfM (Block, 2003). Typical pedagogical tasks (e.g. Information gap or role-plays) from a cognitivist perspective have been considered as cognitive devices (Ellis, 2000) that can elicit the most amount of turns. Despite the different conceptualisations of the term 'negotiation for meaning' (NfM), Block (2003) highlights how the dominant construct of NfM from mainstream cognitivist perspectives has meant that other interpretations have been omitted such as 'negotiation of solidarity and support' (Foster, 1998) and 'negotiation of face' (Davies, 2000). This study conceptualises meaning negotiation as the act of exchanging information between peers (or negotiation for meaning) *and* for different purposes (or negotiation of meaning) following authors (2016). An example of this can be seen in Guth and Helm's (2011) study of a CMC task where learners orally negotiate roles, deadlines and a mode of communication (of meaning) using oral turns (for meaning).

Few studies on task and non-task based spoken interaction in CMC have highlighted external semiotic resources as screen-based, non-embodied modes (not pertaining to humans such as gaze) but include Collentine (2009); Lamy (2006); Lamy (2012); Hampel and Stickler (2012); Liddicoat (2010) and Knight



and Barberà (2016). However, the studies are primarily focused on verbal meaning making. Now we turn to how potential non-verbal meaning making in tasks may be understood through a multimodal perspective.

'Negotiation of Meaning' is also a term used by Wenger as part of Communities of Social practice Theory (Lave and Wenger, 1991). Wenger (1998) describes 'negotiation of meaning' as how we experience the world and our engagement in social processes as meaningful. Participation is an element in the negotiation process that requires active involvement by learners. If all change involves a process of learning, then effective change processes consciously facilitate negotiation of meaning (Wenger, 1998). Learners can engage in legitimate peripheral participation (a viewpoint that makes a fundamental distinction between learning to become part of the community of practice vs intentional instruction). Mediation with tools during participation is imbued with potential meaning not only with regard to what participants learn and do not learn but also what it comes to mean for them (Wenger, 1998). Activities (or tasks) therefore, involving cultural artifacts (technological tools) can 'pass on social inheritance' (Jenlink and Jenlink, 2007, p.381). From this perspective, learners as participants can engage in the process of *learning to be* and thereby construct an identity commensurate with the community of practice (Jenlink and Jenlink, 2007).

A way of understanding and exploring mediated action involving tool-use can be through notions from multimodal (inter)actional analysis (Norris, 2004). In multimodal (inter)actional analysis, different levels of social (mediated) action is a core notion. Higher-level actions and lower-level actions deal with different levels of interaction. Higher-level action is used to refer to large-scale actions,

such as a meeting, and is made up of ‘multiplicity of chained lower-level actions’ (Norris, 2004 p.13). Lower-level action refers to smaller-scale actions, for example, gestures or gaze shifts that become ‘chains of lower-level interactions’ (Norris, 2004 p.13). Lower-level actions support the achievement of the higher-level action.

### **3. General Objectives**

The study aims to understand how meaning is negotiated non-verbally with the navigational screen-based resources in an online task and how that meaning negotiation shapes spoken interaction. These are the research questions:

1. How is meaning made non-verbally with the navigational screen-based resources in an online task?
2. How does the meaning making shape spoken interaction?

### **4. Method**

#### **4.1 Participants**

The participants were adult students in an English as a Foreign Language class as part of their degree programme at a 100% virtual university. The learners stem from two data sets: the first were in a B2.1 level class on the CEFR (upper intermediate) group and engaged in two virtual synchronous peer-peer speaking tasks, one after the other. These were eight participants in the form of four dyads. The second data set were in a level B1.1 class and were engaged in one task with formed by two tasks within it. These were twelve participants also in the form

of six dyads, resulting in ten cases in total. Students were bilingual (Catalan and Spanish) with English as an additional language.

## **4.2 Context**

The participants across the three tasks are presented with various screen-based resources on the interface pages including text and navigational buttons, with a photo in the first task(s) and a timer and pop-ups in the second task. The pages are part of the Tandem tool <http://www.speakapps.eu/#tandem> that facilitates audioconferencing and can distribute different or same materials to different students who are engaged in synchronous spoken interaction tasks.

The first data set was taken from a previous study by Appel, Robbins, Moré and Mullen (2012) in order to explore how different navigation features influenced learning strategies and TL use in tasks. The sampling rationale was to select dyads that demonstrated 'good' performance and followed task design and those who approached the task differently to task design. A result from this study was that only one out of the four cases based on this criteria mentioned the navigational features explicitly and consistently in their talk. Because of the various foci of learners' talk that appeared to take place in relation to the screen, a further study took place (from another classroom, another task type, language level and teacher) to understand how personal and joint meaning making was being carried out across the three tasks using the same tool (authors, 2016). Authors (2016) compared the first data set (info gap and opinion sharing) with a second data set (role-play). There was one interface version (Tandem with navigational buttons) with cases from the second set whereas cases from the first

data set had two versions (html or Tandem with navigational buttons) because of the nature of Appel, Robbins, Moré and Mullen's (2012) study.

Based on the results from the first data set as an exploratory study, the second data set used the same sampling rationale, whereby cases were selected that followed task design and those who appeared to approach the task differently. Specifically, the rationale for case selection were based on findings that learners' relationship with navigational resources affected recursivity of speech, time spent talking in the TL than was designed/expected and that speech could accompany physical moves to navigate (Knight, Barberà and Appel, 2017). Therefore, the selection criteria for case selection for the second data set were the indicators of speech recursivity and the presence/absence of navigational resources in the talk. However, the task type (role-play) appeared to force speech recursivity through learners' need to use questions and answers, so we deemed that this was not a reliable indicator. The focus then turned to explicit oral mentions of the screen-based resources, particularly navigational resources.

Regarding the tasks, cases from the first data set had to collaborate to complete two out of four tasks on the topic of 'travelling'. The first was an Information gap task (spot-the-difference) and the second was an opinion-sharing task, which learners had to navigate to after the first task. Learners had four differences to find in a photo. The second task used the same photo, accompanied by an open question: *What kind of activities can people do in a holiday destination like this?* Students received instructions and guidelines and the tasks were compulsory course assignments. There was no time limit.

The second data set consisted of six cases carrying out a role-play task (divided into sub-tasks) in which learners took turns being the interviewer and interviewee. The first task required that one learner ask questions to his/her peer in their given roles whereas the second sub-task required the interviewer to describe two jobs, one of which their partner had to choose. After the first two tasks are completed (tasks 1 and 2) the two tasks are repeated but the peers change roles (tasks 3 and 4). The task is timed: a timer is present indicating how much time learners have left to do the task and when the time is up, a pop-up (navigational resource) appears. Learners have to click on the pop-up to 'close' before navigating further. Only the first two tasks were analysed, meaning that one student in each case is analysed as interviewer, not both.

#### **4.2.3 Data**

The data consists of recorded oral conversations: the first data set with approximately 23 minutes and the second set with approximately 97 minutes but of which 53 minutes was used (before learners exchanged roles). Recordings were captured by learners using a plug-in for Skype, a free video and audio conferencing tool. Task instructions as a document, tool log files and twenty-two screenshots of the interface that appeared during task (according to task design) were also used.

This current study uses a multimodal and discourse analytic approach in order to gain greater insight on how non-verbal meaning making may be occurring. It revisits the first data set and the second data set (which has become part of a

wider study on how learners mediate turns verbally and non-verbally in a task by authors, forthcoming). This is relevant because the procedure of data analysis forms a part of this wider study. Six cases are presented here because while data saturation is considered to have occurred up until case 4, a closer analysis of case 6 added to what was learnt from analysing and choosing the original four cases.

### **4.3 Procedure**

A qualitative case study approach was used and cases were purposively sampled based on the criteria outlined. In the first data set, one case out of four explicitly, consistently mentioned the navigational resources. Following this, in the second data set, the recordings of all dyads' from one classroom was listened to and initial notes were made about whether learners explicitly mentioned the navigational resources in their talk and which ones. A list of screen-based resources emerged, including navigational resources (from the learner perspective) and was added to iteratively when learners (interviewee or interviewer) mentioned them. This took the form of using phrases like "we can confirm" or learners' replicating the words on the navigational resource in speech e.g. "next task". Incomplete or unintelligible recordings were disregarded. Whilst listening to all the dyads, we made notes about whether learners mentioned the screen-based resources, including navigational resources and which ones from an emic (learner) perspective. A display table was created revealing learners' trajectories of use of the screen-based resources. The table allowed us to compare evidence, identify patterns and see where cases converged and diverged. Conversations of cases chosen followed general transcription

conventions and were segmented according to what was deemed a meaningful utterance and converted to a text document.

In order to attend to the screen-based resources, including navigational resources, the main researcher also carried out researcher task simulation: once with a colleague for the duration of the task and on a second occasion, alone, to focus on the navigational requirements before the task (e.g. to choose the task, to invite peer). Random screenshots of the task were taken during task simulation resulting in twenty-two. This was in order to become familiar with task, minute-by-minute i.e. what, when and how learners should be responding to the screen and compare this with what they were saying in the recordings. Results were triangulated following Yin (2009): comparing evidence from different 1) data sources e.g. tool logs and task instructions (document); 2) perspectives- learner emic perspective and researcher etic perspective plus checks with course coordinator and 3) methods -listening and note taking, transcribing, researcher task simulation and task reconstruction through sequencing screenshots. Regarding the Tandem tool logs, 50% were collected from the six cases in the role-play task because during the data collection of the log files the log files disappeared, as they were not available indefinitely. No Tandem logs were available for the first data set because the identification of the phenomenon had not yet occurred during data collection.

Computer Mediated Discourse analytical perspective (henceforth, CMDA) (Herring, 2004) was used to focus on learner talk. In CMDA the general meaning is studied by focusing on the phenomena of the meaning of words, utterances

(speech acts), macrosegments and raising issues as to what the speaker intends, what is accomplished through language (Herring, 2004). By tracking learners' general trajectory we extend the 'meaning' dimension in CMDA to include how meanings may be being made non-verbally and how navigation can shape oral interaction. Notions from multimodal (inter)actional analysis were also used to understand the relationship between navigational resources, agency and tool-use.

## **5. Results and discussion**

To recap, RQ 1 was 'How is meaning made non-verbally with the screen-based navigational resources in an online task?'

The results we present build on authors (2016) finding that learners have different trajectories of use and non-use of the screen-based navigational resources. Table 1 shows this general trajectory in the role-play task. The ticks indicate explicit oral mention of a navigational resource.



Table 1. Cases' oral mention of screen-based resources in a role-play task including navigational resources.

**Key:**

**Pop-up requiring a response to 'close'**

**'Next Task' navigational resource**

**\*Waiting for confirmation**

**Timer**

Intee: mentioned by interviewee

Inter: mentioned by interviewer

	Waiting for confirmation Task 1	Pop-up Time up!/ close button	Next task	Pop-up Time up!/ close button	Interface page 2 'solution page'	Timer	Next task
Case 1	X	X	√ Inter	√ Intee	x	x	√ Intee
Case 2	X	X	X	X	x	x	X
Case 3	√(before start of task) Inter	√ Inter + Intee	√ Inter	√ Intee	√ Intee	√ Inter	X
Case 4	X	X	X	X	x	x	X
Case 5	x	x	x	x	x	x	x
Case 6	x	x	x	x	x	√ Inter	x

Table 1 shows that cases 1 and 3 (role-play) consistently and explicitly mention the navigational resources. Cases 2, 4, 5 and 6 do not. To compare, we revisit data from the information gap/opinion sharing task as shown in Table 2.

Table 2. Cases' oral mention of navigational resources in an information gap followed by opinion-sharing task.

**Key:**

**Confirmation button after finding a difference**

**'Next Task'**

**See solution** (after opinion sharing task)

	Confirm	Confirm	Confirm	Confirm	Next task	See solution
Case 1 Tandem with buttons	√	√	x	√	√	√
Case 2 Tandem with buttons	√	x	x	x	X	x
Case 3 html versión	x	x	x	x	X	x
Case 4 html version	x	x	x	x	X	x

Table 2 shows that case 1 consistently and explicitly mention the navigational resources. Cases 2, 3 and 4 do not. From this we conclude that some cases allow the tool to sequence/scaffold them during spoken interaction in the TL whereas others do not.

Case 1 (information gap) does not explicitly mention all navigational resources (confirmation 3 is absent) and nor do cases 1 and 3 (role-play). Case 1 (role-play) omits the first two navigational resources (waiting for confirmation and close) and case 3 (role-play) omits the last 'next task'. We conclude that learners that allow the tool to sequence them do not therefore necessarily make all navigational resources explicit everytime they click them. Case 6 (role-play) makes no explicit mention of navigational resources but does mention the time being up ("when the time out we change the roles"). This indicates that case 6 either allowed themselves to be sequenced by the navigational resources but chose to eliminate reference to them while talking or alternatively, chose to time themselves without the screen-based timer. The mention to "time out" is ambiguous as to whether they are referring to the screen-based timer or not. Learners can also start the recording after their partner accepts their invite (as they wait for confirmation) so this may also eliminate this moment in their recordings. It appears that cases that make consistent, explicit mention of navigational resources across the task(s) are allowing the tool to scaffold them even if there are some absences. However, cases also appear to be able to choose not to mention the navigational resources while being sequenced, although this can not be confirmed for case 6. We propose that the cases ability to eliminate reference to the navigational resources is due to their familiarity with the resources (and pop-up) before recording their oral exchange and/or not using the screen resources to sequence them while talking.

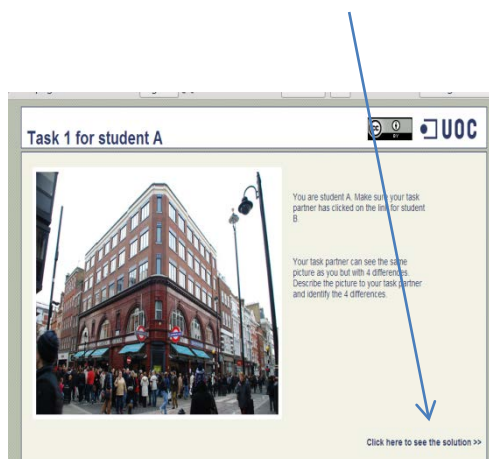
Furthermore, a learner in case 2 (info gap) says "ok. I confirm the one difference. We have the one difference" which is not a word repeated again throughout the

task. This can be explained by learners allowing the navigational resources to sequence them but then choosing to give up. However, this explanation is not reflected in their talk. Another explanation for the reference to the word “confirm” is that case 2 previously allowed themselves to be sequenced (without recording) but carried the word ‘confirm’ over to their recorded interaction (a lexical item taken originally from the navigational resource).

The navigational resources were designed to scaffold learners in the Information-gap task (submit/look at answers and sequence learners as in figure 1) and sequence learners in the role-play task (as shown in figures 2 and 3). These navigational resources are shown below.

#### Navigational resource (Information Gap)

Figure 1. Click to see solution



## Navigational resources (Role-Play)

Figure 2. Time Up! Pop up/

Close

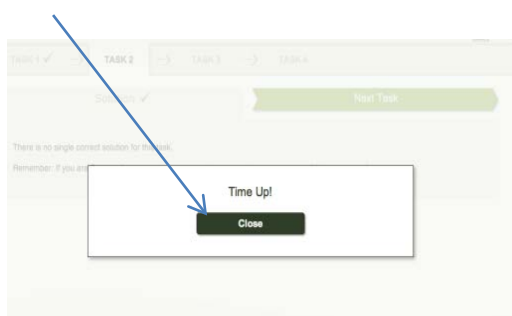
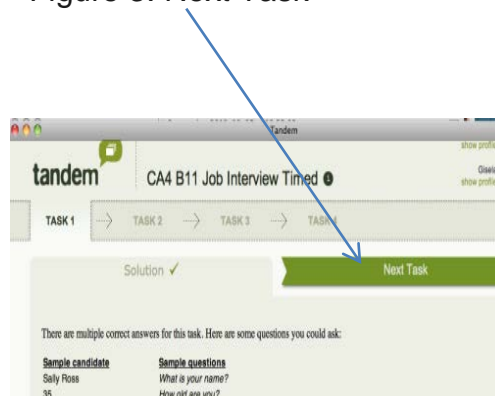


Figure 3. Next Task



Following this, we propose that learners negotiate meaning non-verbally through their experience of the trajectory of use/non-use of navigational resources as they mediate with the tool in the process of task completion.

In the role-play task, the Time Up! pop-up in particular, which would have occurred twice during the two tasks, would have surprised learners, yet no reference or utterance of surprise was made by cases 2, 4, 5 and 6. An explanation for why no explicit reference was made is that learners appear to have already familiarised themselves with the resources (implying navigation) and the tool prior to starting the task. It was possible to enter, navigate through and exit the task (as navigational acts) an unlimited amount of times and which is evident in the Tandem logs (role-play). Being able to exit the tool to prepare or script the spoken interaction was a learner behaviour identified by Appel, Robbins, Moré and Mullen (2012).

The Tandem logs revealed that apart from a general trajectory of allowing the tool to sequence them whilst orally interacting or not, learners also had unique

trajectories, highlighted in figure 4. Whereas Case 1 and 3 were deemed as generally choosing to be sequenced by the tool the frequency and duration that they entered and exited the tool were different and indicated a more specific, nuanced trajectory. Whereas case 1 (role-play) indicates one entry to the tool, cases 2 and 3 (role-play) indicate many more entries. However, case 2 makes more entries lasting over the 11 minutes to do tasks 1 and 2 compared to case 3.

Figure 4. Tandem tool logs for 'The Right Job' Interview (Role-play)

Case 1	Case 2	Case 3																																																												
<p>Tandems</p> <table border="1"> <thead> <tr> <th>Date ▾</th> <th>Exercise</th> <th>Total Minutes</th> </tr> </thead> <tbody> <tr> <td>2015-06-12 10:45:02</td> <td>CA6 What should I do?</td> <td>00:11:58</td> </tr> <tr> <td>2015-05-08 10:07:50</td> <td>CA4: The Right Job Interview</td> <td>00:35:23</td> </tr> </tbody> </table>	Date ▾	Exercise	Total Minutes	2015-06-12 10:45:02	CA6 What should I do?	00:11:58	2015-05-08 10:07:50	CA4: The Right Job Interview	00:35:23	<p>Tandems</p> <table border="1"> <thead> <tr> <th>Date ▲</th> <th>Exercise</th> <th>Total Minutes</th> </tr> </thead> <tbody> <tr> <td>2015-05-08 10:19:41</td> <td>CA4: The Right Job Interview</td> <td>00:40:51</td> </tr> <tr> <td>2015-05-10 11:13:02</td> <td>CA4: The Right Job Interview</td> <td>00:00:04</td> </tr> <tr> <td>2015-05-10 11:30:43</td> <td>CA4: The Right Job Interview</td> <td>00:29:56</td> </tr> <tr> <td>2015-05-10 12:06:32</td> <td>CA4: The Right Job Interview</td> <td>00:24:53</td> </tr> <tr> <td>2015-05-10 12:37:14</td> <td>CA4: The Right Job Interview</td> <td>00:24:12</td> </tr> <tr> <td>2015-05-10 13:14:04</td> <td>CA4: The Right Job Interview</td> <td>00:11:17</td> </tr> <tr> <td>2015-05-10 13:30:35</td> <td>CA4: The Right Job Interview</td> <td>00:04:39</td> </tr> <tr> <td>2015-05-10 16:12:35</td> <td>CA4: The Right Job Interview</td> <td>00:23:34</td> </tr> </tbody> </table>	Date ▲	Exercise	Total Minutes	2015-05-08 10:19:41	CA4: The Right Job Interview	00:40:51	2015-05-10 11:13:02	CA4: The Right Job Interview	00:00:04	2015-05-10 11:30:43	CA4: The Right Job Interview	00:29:56	2015-05-10 12:06:32	CA4: The Right Job Interview	00:24:53	2015-05-10 12:37:14	CA4: The Right Job Interview	00:24:12	2015-05-10 13:14:04	CA4: The Right Job Interview	00:11:17	2015-05-10 13:30:35	CA4: The Right Job Interview	00:04:39	2015-05-10 16:12:35	CA4: The Right Job Interview	00:23:34	<p>Tandems</p> <table border="1"> <thead> <tr> <th>Date ▲</th> <th>Exercise</th> <th>Total Minutes</th> </tr> </thead> <tbody> <tr> <td>2015-05-07 22:21:55</td> <td>CA4: The Right Job Interview</td> <td>00:16:30</td> </tr> <tr> <td>2015-05-07 22:45:11</td> <td>CA4: The Right Job Interview</td> <td>00:00:17</td> </tr> <tr> <td>2015-05-07 22:47:39</td> <td>CA4: The Right Job Interview</td> <td>00:03:55</td> </tr> <tr> <td>2015-05-11 22:14:46</td> <td>CA4: The Right Job Interview</td> <td>00:03:28</td> </tr> <tr> <td>2015-05-11 22:23:55</td> <td>CA4: The Right Job Interview</td> <td>00:18:39</td> </tr> <tr> <td>2015-05-13 22:04:13</td> <td>CA4: The Right Job Interview</td> <td>00:00:00</td> </tr> <tr> <td>2015-05-15 22:36:14</td> <td>CA4: The Right Job Interview</td> <td>00:06:05</td> </tr> </tbody> </table>	Date ▲	Exercise	Total Minutes	2015-05-07 22:21:55	CA4: The Right Job Interview	00:16:30	2015-05-07 22:45:11	CA4: The Right Job Interview	00:00:17	2015-05-07 22:47:39	CA4: The Right Job Interview	00:03:55	2015-05-11 22:14:46	CA4: The Right Job Interview	00:03:28	2015-05-11 22:23:55	CA4: The Right Job Interview	00:18:39	2015-05-13 22:04:13	CA4: The Right Job Interview	00:00:00	2015-05-15 22:36:14	CA4: The Right Job Interview	00:06:05
Date ▾	Exercise	Total Minutes																																																												
2015-06-12 10:45:02	CA6 What should I do?	00:11:58																																																												
2015-05-08 10:07:50	CA4: The Right Job Interview	00:35:23																																																												
Date ▲	Exercise	Total Minutes																																																												
2015-05-08 10:19:41	CA4: The Right Job Interview	00:40:51																																																												
2015-05-10 11:13:02	CA4: The Right Job Interview	00:00:04																																																												
2015-05-10 11:30:43	CA4: The Right Job Interview	00:29:56																																																												
2015-05-10 12:06:32	CA4: The Right Job Interview	00:24:53																																																												
2015-05-10 12:37:14	CA4: The Right Job Interview	00:24:12																																																												
2015-05-10 13:14:04	CA4: The Right Job Interview	00:11:17																																																												
2015-05-10 13:30:35	CA4: The Right Job Interview	00:04:39																																																												
2015-05-10 16:12:35	CA4: The Right Job Interview	00:23:34																																																												
Date ▲	Exercise	Total Minutes																																																												
2015-05-07 22:21:55	CA4: The Right Job Interview	00:16:30																																																												
2015-05-07 22:45:11	CA4: The Right Job Interview	00:00:17																																																												
2015-05-07 22:47:39	CA4: The Right Job Interview	00:03:55																																																												
2015-05-11 22:14:46	CA4: The Right Job Interview	00:03:28																																																												
2015-05-11 22:23:55	CA4: The Right Job Interview	00:18:39																																																												
2015-05-13 22:04:13	CA4: The Right Job Interview	00:00:00																																																												
2015-05-15 22:36:14	CA4: The Right Job Interview	00:06:05																																																												

Although the logs cannot tell us exactly what learners did once they had entered the tool and we cannot assume they were attending to the task, we do consider entering and exiting as an indicator of intentional tool management made up of smaller navigational acts. What is also notable is that cases 2 and 3 (role-play)

managed the tool over a number of days, possibly having a different purpose for each visit. We propose tool (or task) familiarity (Guth and Helm, 2011) as an explanation for numerous visits given that technological tool use can induce anxiety (Appel, Robbins, Moré and Mullen, 2012; Guth and Helm, 2011).

Learners' action/s can be understood both in terms of legitimate periphery participation (Lave and Wenger, 1991) of tool use as well as lower-level and higher-level actions (Norris, 2004). Entering and exiting the tool at different times and over a number of days can be understood as legitimate peripheral participation in the online community. We propose that different behaviours relating to entering and exiting of the tool reflects the means by which learners physically negotiate participation. Whilst some cases did not use the tool as designed, all negotiations of participation require 'active involvement and are legitimate' (Lave and Wenger, 1981).

Learners' use of navigational resources can also be understood as lower-level actions which, when combined together, lead to the higher-level action of technological tool management. In terms of meaning making with the navigational resources, learners' responses to navigational invites to 'close' or as an invite to go to the 'next task' can understood as meaning 'accept' going forward as a completed (computer-human) lower-level action. However, we propose that clicking the same navigational resource going backwards or navigating back and forth overrides the original meaning of the resource as an 'acceptance', and through learners' own navigational sequence, learners re-assign a different meaning to this resource.



Learners appear to have their own 'lines of desire' (Lukin and Du Boulay, 2003) which they carry out using navigational resources. We propose that through non-compliance with the navigational route according to task design, learners negotiate or exhibit a disagreement with it. Tool management appears to be what some cases physically 'negotiate' through their use of navigational resources, before orally negotiating with their peer in the target language (TL).

In order for the cases to complete the pedagogical tasks however, learners needed to have used the navigational resources to access the pedagogical task instructions (text) and thereby entering the tool to get what they needed from the different interface pages. This indicates learner purpose in their navigational acts. Through different trajectories, learners re-purpose the navigational resources: one purpose being to use them to obtain task instructions (using the tool as a kind of repository space) and another as a scaffold to sequence their steps.

Heift's (2002) identification of different personas would be useful to hypothesise that specific learner navigational patterns may resemble the behaviour of 'browsers', 'peekers' or 'adamants'. However, the personas generally related to learners' language proficiency levels, whereas learners in this study are of the same global language level. Proficiency does not explain the different trajectories taken.

Learners alternatively may have found the simultaneous nature of initiating and responding (Coffin and Donohue, 2014) orally and physically, as something that

was cognitively problematic (e.g. Virvou and Tsiriga, 2001) as learners had to attend to the technological interface as well as personal interaction (Liddicoat, 2010). Another explanation is that learners were anxious to carry out synchronous interaction using the tool (Appel, Robbins, Moré and Mullen, 2012) given that they were not familiar with using it. Numerous entries into the tool would have afforded learners tool and/or task familiarity and a means whereby learners could scaffold themselves into how to use it.

With regard to RQ 2: how non-verbal meaning making shapes spoken interaction, we identify two aspects. The first is the impact on the amount of time learners spend interacting orally in the TL. The second is how it shapes their discourse.

Table 3 shows the time required to interact orally in order to complete the role-play task (11 minutes) and the actual time taken. Results appear to confirm that learners chose either to follow the designed navigational path or not. This echoes Knight, Barberà and Appel's finding (2017) that learners' navigational choices can reduce the time spent interacting orally in the TL.

Table 2. Cases and time on tasks 1 and 2 of the role-play

Case	Time required to complete tasks 1 and 2 (Role play)	Time taken by cases (interacting orally)
Case 1	11 minutes	14:33
Case 2		8:07
Case 3		11:48
Case 4		3:33
Case 5		6:20
Case 6		9:31

How the different trajectories impact on learners' talk is also evident in their discourse.

Examples 1, 2 and 3 are excerpts from case 1 (Information Gap). In these examples learners believe they have found a difference and use the navigational resources (buttons) to confirm.

Example 1: Case 1 (Info Gap)

M: this can be the first difference?

F: yes

M: ok I confirm then ok let me (mumbling) ok good then I also see ...

The negotiation that takes place in example 1 consists of first checking with his/her peer, stating who will be the person to confirm (therefore clicking the button), indicating that he is clicking (“let me” then mumbling) and then returning back to the pedagogical task.

Example 2: Case 1 (Info Gap)

F: ok so maybe this is another difference because I don't see that

M: ok we can confirm that. Waiting for confirmation

F. ok and then...

In example 2, a similar pattern emerges with the word ‘confirm’ featuring in the discourse (the same word on the resource) followed by an update to his peer or self-talk saying ‘waiting for confirmation’.

Example 3: Case 1 (Info Gap)

M: I don't know what's the last difference

F: yeah I don't know because your picture is very small. Now I open

M: Ah the colour of sky

F: oh yeah yeah yeah

In example 3, learners are coming to the end of the task and they are not sure about the last difference. F decides to navigate to ‘open’ the answers, we hypothesise to finally know the answers based on his delight/relief expressed in ‘Ahh...’ followed by details of what the last difference is.

Example 4 is also an excerpt from case 1 taken after they have seen the answers and discussed them. They negotiate navigating to the next task using, again, the exact term 'next task' from the navigational resource.

Example 4: Case 1 (Info Gap- Opinion Sharing)

M: ok. Next task?

F: yes ok what is next task? so now we havto discuss...

Results show that the navigational resources provide learners with a source for meta-modal talk, echoing results by Cunningham, Fägersten and Holmsten, (2010) and Fägersten, Holmsten and Cunningham (2010).

In the role-play task, the pop-up is made up of two textual messages: 'Time Up!' and 'Close'. In example 5, the dyads have come to the end of task 1 and the interviewer has asked the questions to the interviewee. The excerpt starts with a response from the interviewee about how much she hopes to earn.

Example 5: Case 3 (Role-play)

L        Oh, well...I have not great aspiration and perhaps a thirty thousand...  
          Euros per year.

A        Okay, perfect! Okay, thank you. I look for the position we have.

A        Okay... well ok.. now... **close** no, I suppose...?

L        I suppose too, I **close** it...

A Well here are the **questions** and the **answers**...We go to the **next task**...?

L Sorry?

A We go to the **next task**?

L Okay, huh...But the computer say me that the **time is up**.

A Ah, yes.

Following task design, learners would have been at the point in the task where the time to complete the task is now up (11 minutes). This is indicated by a timer on the screen changing to red, followed by a pop-up when time has run out. The pop-up enters the discourse, shaping the interaction by becoming a topic of talk. Learners need to attend to the pop-up (visually/textually) as well as physically through clicking the resource (somatically). In this way, meaning appears to be made verbally and non-verbally: multimodally. After closing the button (indicated by "I close it"), the learners negotiate further navigation to Task 2 using the 'Next Task' resource. The words 'next task' becomes part of the utterance.

Another example of how the navigational resources shape the talk is shown in example 6 from case 1 (role-play) where the learners are coming to the end of Task 2 and the interviewee is clarifying some aspects of the two jobs.

Example 6: Case 1 (Role-play)

L Okay, well...

S Do you prefer?

L Yes, I think I prefer the second job. **Time up!** Ok, well I prefer the second job but I don't know if I can choice that because I don't have the experiences in industries but for me it's better, okay?

S Okay good, thank you very much for your time, it's a pleasure er meeting you today (sound interference).

After the interviewee declares "Yes, I prefer the second job", she/he follows with "Time Up! Ok" and continues in the same turn with why she/he prefers the second job. The pop-up interrupts the learner's turn and the learner allows it to become part of his/her utterance but not a topic of talk. Its meaning is not jointly negotiated orally but is negotiated individually verbally (to self) and non-verbally by activating (textual/visual) recognition of it.

Example 7 from case 3 (role-play) is an excerpt from the very beginning of the recording. The learner utters the phrase 'I'm waiting' after 18 seconds.

Example 7: Case 3 (Role-play)

SILENCE

L (after 18 seconds) Mhu

A *Unintelligible* (PAUSE). **I'm waiting.** (PAUSE)

L Well

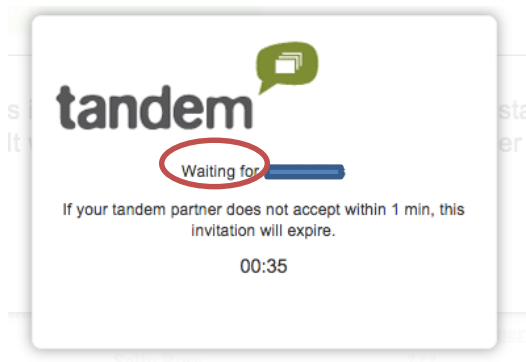
A Okay, (PAUSE) A or two?

L (PAUSE) I have to answer your questions.

A Okay, perfect.

After inviting the partner in the opening of the conversation, following task design a pop-up appears that says 'waiting' as shown in figure 5.

Figure 5. Pop-up after 'inviting' partner technologically



(Learner name deleted)

We hypothesise that the learner's utterance may have been influenced by the text "waiting" on the pop up, as in previous examples. Because a technological response to the technological 'invite' has not been secured we propose that the utterance is either a form of self-talk as learners encounter text, "muttering the



words to himself” (Raudaskoski, 1999, p115) or as a means of updating the teacher of task progression.

An analysis of the discourse reveals a number of things. The first is that navigational acts are sometimes negotiated orally between learners so that they become part of their discourse, confirming results found by Knight and Barberà (2016); Cunningham, Fägersten and Holmsten, (2010) and Fägersten, Holmsten and Cunningham (2010). The exact text on the navigational resources often became embedded in learners’ talk. However, this joint negotiation was consistently absent in cases that did not allow the tool to sequence their steps. The examples also highlight that multimodal communication is taking place both through the screen and with navigational resources on the screen suggesting that meaning making is taking place with two sources of mediational means, namely the screen and their partner and with two main orientations, namely towards their partner and towards the screen. We propose that this scenario represents a multi-party ‘encounter’ (Raudaskoski, 1999) not solely a conversation with their peer. Learners also individually take turns with navigational resources (e.g. ‘accepting’ the computer’s ‘invites’) as found by Liddicoat (2010).

Some cases negotiate meaning orally for pedagogical task completion and orally pertaining to navigational acts, as a form of meta-modal talk. Furthermore, we propose that learners are physically negotiating with navigational resources regarding what to do with specific navigational resources (e.g. next task button) and the tool more generally.

## 6. Conclusions

We have sought to shed light on the non-verbal meaning making pertaining to navigational resources and gain insight as to how they might shape spoken interaction. We have employed a multimodal perspective in order to focus on meaning making and how that is negotiated in a CMC task with navigational resources. Results show that learner-learner oral turns and turn-taking (NfM) is the means for negotiating task content but we also identify that learners physically negotiate with navigational resources. That is to say, learners negotiate their tool-mediated goal-directed action(s), physically through trajectories of tool-use: what learners can/want to do and can not/do not want to do with navigational resources and ultimately the technological tool. This physical negotiation highlights the potential meaning(s) that may be emerging from learners' experience of the navigational resources and the tool. These are not given meanings, but rather negotiated which can also be understood as participation in a community of practice. The application of the notion of lower and higher-level action has given insight into the carrying out of intentional navigational acts as a type of agency, and the means by which learners manage the tool. Learners negotiate how the tool is used and not used (higher level) through the navigational resources (lower level) and as a result, we propose, so are the (higher level) meaning/s that learners may draw from this process. By re-purposing the resources (from sequencing their behaviour to being the means of obtaining what they need) learners reconfigure their roles from tool users to tool managers. We propose that learners are participants engaging in the process of *learning to be* and through the somatic mode (pertaining to the human body) they can negotiate an identity.

That is to say making personal meaning as to who they *are* commensurate with the community of practice. Tool-use may even be considered ‘a contested site’.

Furthermore, navigational acts can be considered a non-verbal way of making meaning that shapes spoken interaction. Navigational acts can form part of the meta-modal discourse. They can also effect the time spent talking in the TL and by implication, the amount of oral turns learners take, which is important for SLA research focused on NfM.

### **6.1 Task design and learning Implications**

This study has focused on navigational resources as mediational tools, however as suggested by Wallace (1999) and echoed by Heift (2002), “the most important mediator of human behaviour in internet environments is the purpose of the people who visit or inhabit them” (p. 5). Indeed, learners’ higher-level purpose/s for tool use seems to emerge in this study as an important design consideration. Therefore, if learners are using/controlling the tool in ways which do not harness the tool’s affordance for a specific pedagogical purpose (in this case synchronous, spontaneous spoken interaction for language learning), then in addition to considering adaptations to the lower-level navigational resources, task/tool design may focus on the higher level action of how learners manage the tool. For example, learner training on how learners can use the tool for different purposes (e.g. planned/scripted spoken interaction vs spontaneous spoken interaction) and the different language learning benefits of each (e.g. less time in the TL vs more time in the TL/ focus on accuracy vs focus on fluency). The tool

design purpose may also need to be rethought in the light of how learners actually use it rather than how it is projected to be used as suggested by Heift (2002).

Although navigational resources as part of the tool may be part of a task designed with the aim of scaffolding learners, learners may use them to scaffold their own purpose(s). They may even re-purpose the tool from one use to the next use.

Therefore, a focus on learner purpose and learner tool experience (experience focus) together with a shift away from the notion of tool affordance (design focus) may serve task designers in understanding learner purposes to improve task processes and outcomes. Whether designers should give learners choice and control over his/her own instruction in order to 'navigate the terrain to compose a unique and individual instructional sequence' (Lawless and Brown, 1997) is perhaps a false dichotomy if learners engage in navigational behaviours that minimise the optimal benefits of a tool so that little language learning gains are made. Focusing on learners' experience and purposes might help to reframe the issue. As Mantovani (1996) suggests, CMC participants cannot be regarded simply as technology users but rather, they are social actors with their own aims and autonomy in situations and it is technology that must adapt to them.

## References

- Ahearn, L. M. (2001). Language and agency. *Annual review of anthropology*, 30(1), 109-137.
- Amaral, L. A., & Meurers, D. (2011). On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. *ReCALL*, 23(01), 4-24.
- Appel, Robbins, Moré, & Mullen, (2012). Task and Tool Interface Design for L2 Speaking Interaction Online. In L. Bradley & S. Thouësny (Eds.), *CALL: Using, Learning, Knowing*, EUROCALL Conference, Gothenburg, Sweden, 22-25 August 2012, Proceedings (pp. 15-19). Dublin: Research-publishing.net, 15.doi: [10.14705/rpnet.2012.9781908416032](https://doi.org/10.14705/rpnet.2012.9781908416032).9781908416032
- Blake, R. (2016). Technology and the four skills, *Language Learning & Technology*, 20(2), 129-142.
- Blin, F., & Jalkanen, J. (2014). Designing for language learning: agency and languaging in hybrid environments. *APPLES: journal of applied language studies*, 8(1), 147-170.
- Block, D. (2003). *The social turn in second language acquisition*. Washington, DC: Georgetown University Press.
- Block, D. (2013). Moving beyond Lingualism: Multilingual Embodiment and Multimodality in SLA. In S. May, (Ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 54-77). New York, NY and London, UK: Routledge.
- Bruner, J. (1987). *The role of dialogue in language acquisition*. In S. R. J. Jarvella & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 214-256). New York: Max-Plank-Institut for Psycholinguistik.
- Calvo-Ferrer, J.R., Melchor-Couto, S. & Jauregui, K. (2016). Multimodality in CALL. *ReCALL*, 28(3), 247-252.
- Coffin, C., & Donohue, J. (2014). A language as social semiotic-based approach to teaching and learning in Higher Education, *Language Learning*, 64, Supplement 1, 287-297.
- Collentine, K. (2009). Learner use of holistic language units in multimodal, task-based synchronous computer-mediated communication. *Language Learning & Technology*, 13, 68–87.

- Cunningham, U. M., Fägersten, K. B., & Holmsten, E. (2010). "Can you hear me, Hanoi?" Compensatory mechanisms employed in synchronous net-based English language learning. *The International Review of Research in Open and Distributed Learning*, 11(1), 161-177.
- Davies, J. (2000). *Social Perspectives on negotiation for meaning*, unpublished MA dissertation, Institute of Education, University of London.
- Ellis, (R). (2000). Task-based research and language pedagogy. *Language Teaching Research* 4(3),193–220.
- Fägersten, K. B., Holmsten, E., & Cunningham, U. (2010). Multimodal communication and meta-modal discourse. In R. Taiwo (Ed.), *The handbook of research on discourse behavior and digital communication: language structures and social interaction* (pp. 144-163). Hershey, PA & New York: Information Science Reference.
- Ferretti, F., Adornetti, I., Cosentino, E., Marini, A., (2013). Keeping the route and speaking coherently: the hidden link between spatial navigation and discourse processing. *Journal of Neurolinguistics*, 26(2), 327–334.
- Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied linguistics*, 26(3), 402-430.
- Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, Mass.: MIT Press.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom* (Vol. 428). Portsmouth, NH: Heinemann.
- Guth, S., & Helm, F. (2011). Developing multiliteracies in ELT through telecollaboration. *ELT journal*, 66(1), 42-51.
- Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(02), 116-137.
- Heift, T. (2002). Learner control and error correction in ICALL: Browsers, peekers, and adamantans. *CALICO Journal*, 19(2), 295–313.
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338-376). New York: Cambridge University Press.
- Jenlink, P. M., & Jenlink, K. E. (2007). Implications of Cultural Psychology for Guiding Educational Practice: Teaching and Learning as Cultural Practices. In J. L. Kincheloe, & R. A. Horn Jr., (Eds.), *The Praeger*

*Handbook of Education and Psychology*, Volumes 1-4 (pp. 374-384), Connecticut: Praeger Press.

- Knight, J., & Barbera, E. (2016). The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks. In Pixel (Ed.) *Proceedings of the 9th International Conference on ICT for Language Learning* (pp. 248-252). Italy: Libreriauniversitaria.it
- Knight, J., Barbera, E., & Appel, C. (2017). A framework for learner agency in online spoken interaction tasks. *ReCALL*, 29(3), 276-293.
- Knight, J., Dooly, M. & Barberà, E.(2018). Navigating a multimodal ensemble: learners mediating verbally and non-verbally on online interaction. Manuscript submitted for publication.
- Lamy, M-N. (2006). Multimodality in second language conversations online: Looking for a methodology, In A. Baldry & E. Montagna (eds.), *Interdisciplinary perspectives on multimodality: Theory and practice*. Proceedings of Third International Conference on Multimodality, (pp. 237-254). Campobasso: Palladino
- Lamy, M. N. (2012). Click if you want to speak: Reframing CA for research into multimodal conversations in online learning. *International Journal of Virtual and Personal Learning Environments (IJVPLE)*, 3(1), 1-18.
- Lantolf, J. P. (ed.). (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge university press.
- Lave, J., & Wenger, E. (2002). Legitimate peripheral participation in communities of practice. In M. Lea, & K. Nicoll (Eds.), *Distributed learning: Social and cultural approaches to practice* (pp. 56-63). London: Routledge.
- Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional science*, 25(2), 117-131.
- Levy, M., & Kennedy, C. (2004). A task-cycling pedagogy using audio-conferencing and stimulated reflection for foreign language learning, *Language Learning and Technology*, 8(2), 50-68.
- Liddicoat, A. J. (2010). Enacting participation: Hybrid modalities in on-line video conversation. In C. Develotte, R. Kern, & M. N. Lamy (eds.), *Décrire la conversation en ligne*. (pp. 37–50). Lyon, France: ENS Éditions.
- Luckin, R., & Du Boulay, B. (2003). Lines of desire: the challenges of interactive educational TV. *Proceedings of the 11th International conference on*

*artificial intelligence in education (AIEd 2003)* <http://rp-www.cs.usyd.edu.au/~aied/papers.html>

- Mahdi, H.S. (2014). The impact of computer mediated communication environments on foreign language learning: A review of the literature, *Teaching English with Technology*, 14(2), 67-86.
- Martin, J. (2004). Self-regulated learning, social cognitive theory, and agency, *Educational psychologist*, 39(2), 135-145.
- Mantovani, G. (1996). *New communication environments: from everyday to virtual*. London: Taylor & Francis.
- Mercer, S. (2011). Understanding learner agency as a complex dynamic system, *System*, 39(4), 427-436.
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. New York: Routledge.
- Novick, D. G & Sutton, S. (1997). What is mixed-initiative interaction. In the working notes of the AAI spring symposium on computational models for mixed initiative interaction (pp. 114-116). AAI/MIT Press.
- Raudaskoski, P. (1999). *The use of communicative resources in language technology environments. A conversation analytic approach to semiosis at computer media* (Doctoral dissertation). Retrieved from <http://vbn.aau.dk/ws/files/72280035/pirkkosphd.pdf>
- Riva, G. (2001). Communicating in CMC: making order out of miscommunication. In L. Anolli, R. Ciceri, & G. Riva (eds.), *Say Not to Say: New Perspectives on Miscommunication* (pp. 197-225). Amsterdam: IOS Press.
- Van Leeuwen, T. (2004). *Introducing Social Semiotics: An Introductory Textbook*. London: Routledge.
- Van Lier, L. (2008). Agency in the classroom. In J. P. Lantolf & M. E. Poehner (eds.), *Sociocultural theory and the teaching of second languages* (pp. 163– 86). London: Equinox.
- Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied linguistics*, 6(1), 71-90.
- Virvou, M. and Tsiriga, V. (2001). Web Passive Voice Tutor: an Intelligent Computer Assisted Language Learning System over the WWW. In Okamoto, T., Hartley, R., Kinshuk, and Klus, J. (Eds.) *IEEE International Conference on Advanced Learning Technologies: Issues, Achievements and Challenges*, (pp. 131-134), Los Alamitos, CA: IEEE Computer Society Press.



- Vygotsky, L. S. (1981). The instrumental method in psychology. In J. V Wertsch (ed.), *The concept of activity in Soviet psychology* (pp. 134-143), Armonk, NY: Sharpe.
- Wallace, P. (1999). *The psychology of the Internet*. New York: Cambridge University Press.
- Wenger, E., (1998). *Communities of Practice: Learning, Meaning, and Identity*, Cambridge: Cambridge University Press.
- Yin, R. K. (2009). *Case study research: Design and methods (4th ed.)*. Thousand Oaks, CA: SAGE.

17.5 Paper 6: But the computer say me the time is up: the shaping of oral turns mediated with and through the screen

**But the computer say me the time is up**

**But the computer say me the time is up: the shaping of oral  
turns mediated *with* and *through* the screen**

**But the computer say me the time is up: the shaping of oral turns mediated *with* and *through* the screen**

**Abstract**

Research on mode in computer-mediated communication and language learning has primarily focused on mode-as-channel of communication such as audioconferencing. However, technological tools that facilitate communication are becoming more sophisticated and diverse, increasing the multiple ways in which learners can convey and respond to meanings, both orally/visually through the screen, when interacting with their peers and with screen-based resources. Given that the role of mode is under researched in task-based language learning (Gilbert, Manchón and Vasylets, 2016) such research may offer teachers, designers and learners insights for understanding how meaning is made with both verbal and non-verbal screen-based resources as modes. In response, this paper explores how spoken interaction is shaped by screen-based resources during an online audioconferencing task. Audio recordings of six dyads are analysed using discourse analysis alongside interface screenshots with a focus on turns and turn-taking between peers and screen-based resources.

Results reveal that screen-based resources can become: 1) embedded or modified in oral turns; 2) resources to initiate and support oral turns; 3) topics of talk whereby learners exchange oral messages about procedural aspects, task status and the internal state of the learner (Kraut, Fussell, and Siegel, (2003)). The study also reveals that learners can orientate towards some resources as 4) turn-takers, such as pop-ups, as if these non-verbal resources were participants in the interaction. Learners can initiate and/or respond to resources orally and/or somatically revealing how occasionally peer-to-peer talk can resemble a multi-party, multimodal encounter (individuals and non-human resources). Resources as active or passive agents in physical turn-taking during oral CMC tasks are also highlighted.

**Key words:** computer-mediated communication, multimodal turn-taking, screen-based resources, spoken interaction, task based language learning, semiotic resources

## **1. Introduction**

Screen-based visual elements play an increasingly more prominent role in modern society (Liou, 2011) such as interactive banking, getting medical appointments etc. and which can now be considered the norm. This is also becoming evident in online education. Online language learners carrying out spoken interaction tasks can use audioconferencing or videoconferencing tools or systems for computer-mediated communication (henceforth CMC). During oral interaction learners face the possibility of a 'multimodal ensemble' (Bezemer and Kress, 2016) on the screen. In any session, learners may be attending to the semiotic resources emanating from their peer (e.g. aural inputs, gesture and gaze) and/or semiotic resources from the screen (screen-based resources) (e.g. text, visuals, audio). These may also include screen-based navigational resources that learners need to engage with physically through clicking. In short, the interplay of potential resources in an online language learning environment is highly complex and requires considerable effort on behalf of the learner.

Therefore, in order to understand the complexity of all the modes involved in meaning creation (Calvo Ferrer, Melchor-Couto and Jauregi, 2016), such as the scenario described, both learner processes and outcomes of spoken interaction tasks seem increasingly in need of a multimodal lens: a lens that not only considers verbal and non-verbal (textual, visual, tactile/somatic) communication but also embodied and non-embodied communication (emanating from humans *and* from computers). However, not only is there a lack of research on the impact of multimodal communication in online language classrooms (Hampel and Stickler, 2012) but also the role of mode is under researched in Task-based learning research (Gilbert, Manchón, Vasylets, 2016). Furthermore, within the field of second language acquisition (henceforth SLA), the predominant way to analyse such a scenario has been by focusing on learners' meaning making through purely verbal means, highlighting the 'lingual bias' (Block, 2013). It follows, therefore, that a shift in how meaning making is both conceptualised and analysed might take into account the screen that is present and the learners' potential to initiate and/or respond, both orally and physically, to screen-based resources while interacting with their peers, thus moving away from a language-only focus. Most studies in SLA have predominantly focused on mediation using spoken language as a semiotic resource *through* the screen by focusing on different channels of communication (or bi and tri-channels) and its/their impact on human-human spoken interaction. This study focuses on mediation with spoken language as a semiotic resource *through* the screen as well as the external semiotic resources *on* the screen.

## **2. Theoretical Framework**

### ***2.1. Computer Mediated Communication for Spoken Interaction***

Computer Mediated Communication (CMC) tools offer practice for speaking and human spoken feedback opportunities either with other L2 learners or L1 speakers (Blake, 2016). Synchronous CMC tools can facilitate videoconferencing (e.g., Adobe Connect, Big Blue Button, Blackboard Collaborate, Skype, Google Hangout, Zoom) and typically allow learners to exchange video, images, and text in real time and are important tools for developing L2 speaking practice (Blake, 2016). In addition, audioconferencing and audiographic conferencing tools may offer the same affordances but without the added presence of the teacher/s or classmate/s face/s and body. Some tools (e.g. *netmeeting*, *Tandem* from SpeakApps project) and computer systems (e.g. *Lyceum*) allow for audioconferencing to be a channel with 'voice, but no image' (Yamada, 2009). However, learners can indeed face screen-based (non-embodied) image and/or other content with audioconferencing tools. This 'input' may include content from an interface(s) pertaining to (inter)net-based or non net-based sources.

#### ***2.1.1 Screen-based resources and turn-taking in audio and video CMC for spoken interaction***

Levy and Kennedy's study (2004) highlighted the possibilities for screen sharing using the audioconferencing tool *netmeeting*. Learners could share the same screen while browsing the net and talking simultaneously. Non networked-based audioconferencing tools, such as the *Tandem* tool (which this current study focuses on), facilitated the distribution of different (or same) screens to learners so that screen-based textual instructions and navigational resources may be the same or different, as found in Knight and Barberà's study (2016). In videoconferencing, screens can be minimized, meaning that the interlocutor's face can 'disappear'. Learners hear sound from the video (as in audioconferencing) but can in fact see net-based content (internet pages). An example

of this is outlined in Guth and Helm's (2012) study that used Skype audio and video for telecollaboration.

Different studies on CMC for spoken interaction, focusing on teaching and language learning, have used many methodologies and have taken different foci of study. These include statistical analysis, a social semiotic approach, case study, discourse analysis and conversational analysis (see Jenks, 2014 for an overview of methodologies, studies and authors). These studies are predominantly concerned with spoken interaction, including turn-taking with speech as the mode *through* the screen.

However, because this current study is focused on learners' mediation (including possible turn-taking) *with* the screen through the use of various modes while simultaneously carrying out oral interaction with their peers *through* the screen, we draw on studies from the field of SLA and CMC (audio and video). We take into consideration CMC studies that are task-based and non-task based as well as some pertinent studies from the literature on Human-Computer Interaction (henceforth HCI). This is in order to gain insight into how oral turns may be shaped by screen-based resources while participants interact orally, as well as to gain insight into the non-verbal turns that may be occurring with screen-based resources.

In SLA, some refer to audioconferencing as 'voice, but no image' (Yamada, 2009). However, a number of the above studies have highlighted the screen-based images and texts as well as other resources (or modes) that do not necessarily pertain to another person (e.g. gaze and gesture) while speaking. Lamy's (2006) study on an audiographic conferencing system focused on turn-taking and identified the semiotic resources for spoken interaction to include natural language (written and spoken), visual resources such as icons, images, colours and shapes. She also found that the notion of sequentiality in conversation analysis could be used to characterise conversational moves within electronically-mediated multimodal conversations, albeit redefined: when more than one modality was available, turns were managed in a hybrid mode across speech and text so that turns were constructed multimodally (spoken and written turns), therefore operating transmodally (across modes).

Knight and Barberà's study (2016) on joint and negotiated meaning construction using a non net-based audioconferencing tool (*Tandem*). The tool distributed different interface (screens) to learners, highlighting how language learners were multi-tasking in spoken language tasks using different screen-based resources (spoken language, a photo, navigational resources) to achieve different smaller purposes within a task. Following this, authors' study (forthcoming) found learners took oral turns with peers *and* with the various screen-based resources (textual instructions, pop-ups, navigational resources) during a spoken interaction task, highlighting how turns could be shared multimodally with screen-based resources. The two studies underscored how learners were required to be both 'semiotic initiators and responders' (Coffin and Donohue, 2014): initiating and responding to (oral/aural) meaning with their peer as well as potential (textual/visual/navigational) meaning(s) on the screen. The studies also highlighted that while audioconferencing can be considered to be a monomodal channel of communication (audio) as Yamada (2009) suggests, the screen can also serve to frame many different external modes-as-semiotic-resources.

In a study on videoconferencing, Liddicoat (2010) focused on human-human interaction in CMC using videoconferencing and highlighted the turns made through human-to-human interaction as well as turns made at specific points through human-computer interaction. He noted the non-embodied external resources used by learners in turn-construction. At the beginning of an online interaction the initiator of the conversation had to firstly secure the attendance of a non-present co-participant. This which was achieved through technology, namely a message via the computer (screen): 'Andrew wants to have a video conversation' which is neither spoken or written by Andrew but is initiated by him with technology. The response by his partner is either a choice to press 'respond' or 'refuse'. This sequence, according to Liddicoat (2010), resembled summons-answers sequences where the verbal equivalent may be 'hey' or naming; and the technological equivalent may be the ringing of the telephone. He also noted that the online conversation involved a multimodal accomplishment of openings, interruptions and closings which, following Tudini (2014, p. 3), "needed to be both technologically and interactionally accomplished due to constraints of the medium."

In the HCI and CMC literature, a number of studies have explored video mediated and audio-mediated communication. Studies have focused on how the affordances of CMC affects conversational participation, including how the carrying out of non-language learning (physical) tasks shape oral utterances. (e.g. Clark and Brennan, 1991; Dourish, Bellotti, Mackay and Ma, 1993; Sellen, 1995; Kraut, Fussell, and Siegel, 2003).

Dourish, Belotti, Mackay and Ma (1993), noted the affordances of some electronic systems to use daemons (active agents) that seek out information of interest to a user and generate "reminder" events. These may include the generation of synthesised speech or non-speech audio cues through the media space. Kraut, Fussell and Siegel (2003) considered the ways in which visual information was used as a conversational resource in the accomplishment of collaborative physical, (non-language learning) tasks and the effects of this visual information on performance and on conversational strategies. They noted that although most systems (e.g. videoconferencing, electronic mail, audioconferencing) are designed to support group activities that can be performed without reference to the external spatial environment (e.g. decision making) they highlight that observational studies of physical collaboration suggest that people's speech and actions are intricately related. In particular, to the position and dynamics of objects, other people, and ongoing activities in the environment. Conversations during collaborative physical tasks typically focus on the identification of target objects, descriptions of actions to be performed on those targets, and confirmation that the actions have been performed successfully. Kraut, Fussell and Siegel's (2003) comparative study of collaborative bike repairing using video and audioconferencing systems and the utterances that occur in the process, relate to different message types. *Procedural* (instructions furthering task completion), *Task status* (state of task or objects in the task), *Referential* (pertaining to the identifications or location of task objects), *Internal state* (Intentions, knowledge, emotions etc.).

Up until now our literature review for CMC for spoken interaction has focused predominantly on CMC and task-based SLA for spoken interaction, non-language learning CMC studies in which screen-based resources appear to form turns and pertinent studies from the HCI literature. We now expand our review further to focus on non-verbal, turn-taking with screen-based resources and studies that have considered screen-based resources beyond the lingual. This is in order to form a framework for analysing verbal and non-verbal discourse in relation to turns and turn-taking.

## ***2.2 A framework for analysing verbal and non-verbal turns and turn-taking***

The analysis of discourse in CMC has often been carried out using “traditional” tools such as discourse analysis or conversational analysis (Benson, 2015). Although historically, spoken interaction has been a central focus through the work of two major schools, namely conversational analysis (CA) (Sacks et al., 1974) and the Birmingham school of discourse analysis (Sinclair and Coulthard, 1975), some have queried how far sets of principles from these schools could be applied to spoken interaction through CMC. However, CMC researchers have largely relied on these principles often adapting them (Benson, 2015). Benson (2015) notes that although the various forms of multimodal CMC (e.g. blogs and social media) clearly differ from spoken interaction (e.g. sometimes only being ‘sporadically conversational’ as found by Herring, 2005), the degree to which they resemble it is an important issue. This leads to the question as to what steps need to be taken in order for these analytical tools to be used for other multimodal discourses (Benson, 2015).

In this study, we also carry out steps so that tools for analysis of discourse can be applied to screen-based modes and resources alongside (but not solely pertaining to) spoken language. We build on Benson’s (2015) approach of using discourse categories of ‘exchange’, ‘turn’, ‘move’ and ‘act’ (Coulthard 1985; Stenström, 1994) that he applied to the multimodal processes that generated a YouTube page, including written comments but also a series of language learning videos that when analysed together as a whole, Benson considered a CMC ‘text’. Importantly, Stenström inserts the category of ‘turn’ (Sacks et al., 1974) between the category of ‘exchange’ (turns are part of an exchange) and ‘move’ in the categories for discourse analysis, which is pertinent to the focus of this study. In Benson’s (2015) study, a turn is a basic unit of interactional analysis and turns are framed by the affordances of the interface page, which governs how users’ contributions will appear on the page. If a user contributed semantic content it counted as a turn (e.g. writing a comment and uploading it or clicking a ‘like’ action). According to Benson (2015) “in order for a turn to be interactive, this action, whatever its form, must somehow be linked to an action performed by another user” (2015, p. 88).

Furthermore, because turn-taking with a minimal initiation/response is one way that ‘an encounter’ (Raudaskoski, 1999) may occur, leading to semiotic mediation or semiosis, the main focus of analysis is on the interpretative work of learners in their ‘encounters with others’ (Raudaskoski, 1999). Raudaskoski (1999) identifies how the ‘other’ has been generally approached as 1) text or discourse, 2) a social entity and 3) a sign. These three conceptualisations are used to understand the data.

The first conceptualisation of the ‘other’ as text, is one way in which text/discourse can be approached by focusing on language as action “by somebody to do various things in the world (language understood as a present event, with focus on the effects and interaction)” (Raudaskoski, 1999, p. 40). Speech act theory (Austin, 1962; Searle, 1969 and Grice, 1975) is part of this approach and is based on speakers’ intentions (realised as illocutionary acts) such as promising, ordering, greeting or instructing.

The second general approach sees the ‘other’ as a social entity. This usually implies agency where the whole of society works through its agents’ interactions with each other (Raudaskoski, 1999). Therefore, any encounter can be regarded as two or more social agents



meeting. Raudaskoski (1999) offers an example of this whereby the computer is referred to as “this one” as in “this one speaks American too”, as another social entity. This approach has the possibility to analyse human-computer dialogue. However, according to Raudaskoski (1999), the spatio-temporal meaning construction is one-sided, i.e. the human user is responsible for constructing the meaning on the basis of the designer’s represented turns.

The third approach conceptualises ‘the other’ as sign and is researched in the field of semiotics. The relation between signs and meaning has been the main focus of semioticians. “This has meant an emphasis on how the object, sign and interpretant are interrelated in the mind of the receiver of the sign” (Raudaskoski, 1999, p. 58). Semiotic materiality proposes that no sign, be it a verbal or nonverbal one, has only one stable meaning, but that the complexities of the communicative situation influence semiosis. “A sign can have various interpretants, different meanings, instead of being semiotically consistent” (Raudaskoski, 1999, p. 58).

Finally, two main underlying principles support the analysis of turns and turn-taking with respect to potential screen-based ‘others’ in the framework. The first is that of ‘relevance’, which echoes the premise that “modes do not exist without social actors utilising them in some way” (Jewitt, Bezemer and O’Halloran, 2016, p. 115) and the second is ‘orientation’. Benson (2015) employed the notion of ‘orientation’ from Conversational Analysis (henceforth CA) to understand moves that he conceptualised as ‘responses’ on a YouTube page (e.g. a video response, ‘like/’dislike’ icons or written ‘comments’) to an ‘initiation’ (uploading a video). From discourse, Sacks et al. (1974) a founder of CA, proposes that by looking at subsequent turns a researcher can ground an analysis of previous ones. For example, an identification of an “accusation” can be grounded by observing how the recipient responds to it (e.g. with an excuse or justification). Therefore, by focusing on learners’ orientation to the prior turn we can see what a turn’s talk is occupied with or what the learners’ understandings of the previous turn is. Norris (2006, p. 4), refers to a similar notion of “how others react to the messages”, from a multimodal perspective.

### **3.Purpose of Study and Research Questions**

The purpose of this study is to understand what and how oral turns are shaped by learners’ encounters with screen-based resources (by focusing on how this is made relevant in their oral turns). In addition, the study aims at understanding how learners may initiate or respond through non-verbal semiotic and screen-based resources while they interact orally with their peer.

#### **RQ 1**

How are learners’ oral turns shaped by encounters with the available screen-based resources in an online task?

#### **RQ 2**

(How) Do learners initiate or respond to other screen-based resources and if so, what does this tell us about turn-taking where screen-based resources are considered as potential ‘others’?

### **4. Participants and Context**

The participants were students in an English as a Foreign Language class, part of their degree programme at a 100% virtual university. The learners were in a B2.1 level on the CEFR (upper

intermediate). There were 12 adult students: 2 male and 10 female, 26-55 years old, engaged in a virtual synchronous peer-to-peer speaking role-play task. Students were bilingual (Catalan and Spanish) with English as an additional language. The students form 6 dyads as case studies. Data sources include approximately 97 minutes of peer-to-peer recorded oral conversations, of which 53 minutes were transcribed in broad transcription (before learners changed roles to repeat the task). In addition, screenshots of the task were taken randomly by the main researcher (samples of these can be seen in Appendix A).

The task was a role-play activity in which learners had to take turns being the interviewer and the interviewee. Only audio was available, so they could not see each other. The role-play task was sub-divided into two tasks (Task 1 and Task 2, which was then repeated as learners changed roles). Task 1 required the learners to ask questions in the role of interviewer and the peer to respond as the interviewee, as can be seen in the screenshot (figure 1). Interviewers were required to ask minimally 5 questions and prompted to elicit some information from the interviewees. Sample textual screen-based information was provided. In Task 2 (see figure 2), interviewers were instructed to describe two jobs to their partner. Interviewees were instructed to ask 3 questions about the jobs and say which they preferred. The tasks were timed with a timer showing how much time learners had left and a pop-up emerged when the time was up.

Figure 1. Screenshot of Task 1 in the role of interviewer

The screenshot shows a language learning interface for a job interview role-play task. At the top, there is a progress bar with 'TASK 1' selected, followed by 'TASK 2', 'TASK 3', and 'TASK 4'. Below the progress bar is a green button labeled 'See Solution'. A timer at the top right shows '02:35'. The main content area is divided into two columns. The left column contains instructions for the role-play, and the right column contains sample information and questions.

**Instructions:**

This activity is a role play where Student A will pretend to be a **job consultant (the interviewer)** and Student B will be a **job applicant (the interviewee)**.

In this first task, **Student A (the interviewer)** needs to find out some information about Student B (the interviewee).

Time limit: **4 minutes**.

**ATTENTION: You and your partner have exactly 4 minutes to complete this task. Use your time efficiently!**

**Sample information:**

Here is some sample information from a previous job applicant. Use it to create questions and find out some important information about Student B. (Two sample questions are provided.)

**Ask your partner at least 5 questions.**

Sample candidate	Your partner (sample questions)
Sally Ross	???
35	???
Boston	Where do you live?
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible. Friendly.	What words describe you?
Driver's license	???
40,000€ a year	???
9 am - 5 pm	???

Current Call

Figure 2. Screenshot of Task 2 in the role of the Interviewer.

The screenshot shows a task interface with a progress bar at the top indicating 'TASK 1' is completed and 'TASK 2' is active. A 'See Solution' button is visible. Below the progress bar, a timer shows 'TIEMPO RESTANTE DE LA TAREA' at 06:31. The main content area contains instructions for two students and two job descriptions.

**Student A (the job consultant) has two possible jobs available. S/he should explain the jobs to Student B (the candidate).**

**Student B should ask three questions about the two jobs and then say which job s/he prefers and why.**

**Time limit: 7 minutes**

**Here are two different jobs you have available. Briefly describe both of them to your partner. Then answer your partner's questions about the jobs. (If you don't know the answers, be creative!)**

<b>Job 1: Bilingual Sales Manager</b>	<b>Job 2: Travel agent</b>
London, England, New York, NY, or Madrid, Spain	Any major city in Europe
Must be willing to travel to other cities once a month	Must be willing to travel to other cities once a month
High school diploma required	High school diploma and business school required
Spanish and English required	English required; a foreign language (Spanish, French or Chinese) is highly desired
10+ years of sales experience required	2 years experience required in the travel industry or in business
Microsoft Word, Excel, Powerpoint and Access required	Microsoft Word, Excel, and Powerpoint required
Must be a good salesperson	Must be a good communicator
9 a.m. - 5 p.m. or 12 - 8 p.m.	9 a.m. - 6 p.m.
Must have a car	No car required
€30,000 a year + commission	30,000€ a year

After the two tasks were completed (indicated by the banner 'TASK 1' and 'TASK 2' in figures 1 and 2) the tasks were repeated but the peers changed roles (indicated by the banners 'TASK 3' and 'TASK 4' in figures 1 and 2). Learners had to navigate from one task to the next and tasks were timed.

## 5. Instruments and Analysis Procedure

For the analytical procedure, the main researcher compared the transcripts with the screenshots to check whether there was coherence between what learners were saying and the natural progression of the task or not. To answer research question 1, where words or phrases in the transcripts exactly or partially matched (e.g. 'agent' in text becomes 'agency' in spoken language), the resources in the screenshots were then highlighted in bold text in the transcripts. We identified and noted if and how each case was encountering and using the text in some way and compared this across the cases. In the analysis, we looked for evidence as to when and how learners were using screen-based resources in their speech. In this process, we developed the categories iteratively and could then identify what learners did with and/or how they used screen-based resources in relation to the content of their turns (the 'what' of the turns). We also looked at how learners orientated (following CA principles) to screen-based resources (the 'what/who' of the turn-taking).

Only the first two tasks were analysed, focusing on one learner as initiator/responder in the two tasks presented.

## **6.Results and Discussion**

To answer RQ1 we identified that learners encountered and mediated with speech as a mode (spoken language as a resource) and other screen-based modes and resources throughout the task. This shaped the spoken turns and turn-taking between the dyads in different ways.

In relation to speech as mode used for interaction with their peer, an initiation/response sequence is evident through learners' verbal turns, many in the form of questions and answers e.g. in example 1 (lines 1 and 2), example 2 (lines 4 and 5) and example 3 (lines 10 and 11). Questions and answers were the most frequent turns and this is perhaps not surprising given that the task is to carry out an interview where question/answer pairs are the social norm. All names have been anonymised.

Example 1 (case 2): question and answer

- 1 N Hello, good afternoon, (pause) how are you?
- 2 G I'm fine, and you?
- 3 N fine thank you

Example 2 (case 1): question and answer

- SILENCE and rustling of paper
- 1 L (after 11 seconds) Ok paper  
SILENCE and rustling of paper
  - 2 S (after 23 seconds) Hi Laia
  - 3 L Hi
  - 4 S How are you today?
  - 5 L I'm fine thanks
  - 6 S Erm Sorry, what is your name?
  - 7 L My name is Laia, Laia Baños
  - 8 S Okay...
  - 9 L Yeah
  - 10 S sorry... how old are you?
  - 11 L I'm 34 years old
  - 12 S Er Well, Where do you live?
  - 13 L I'm live...I live.. I'm living near *of* Girona.

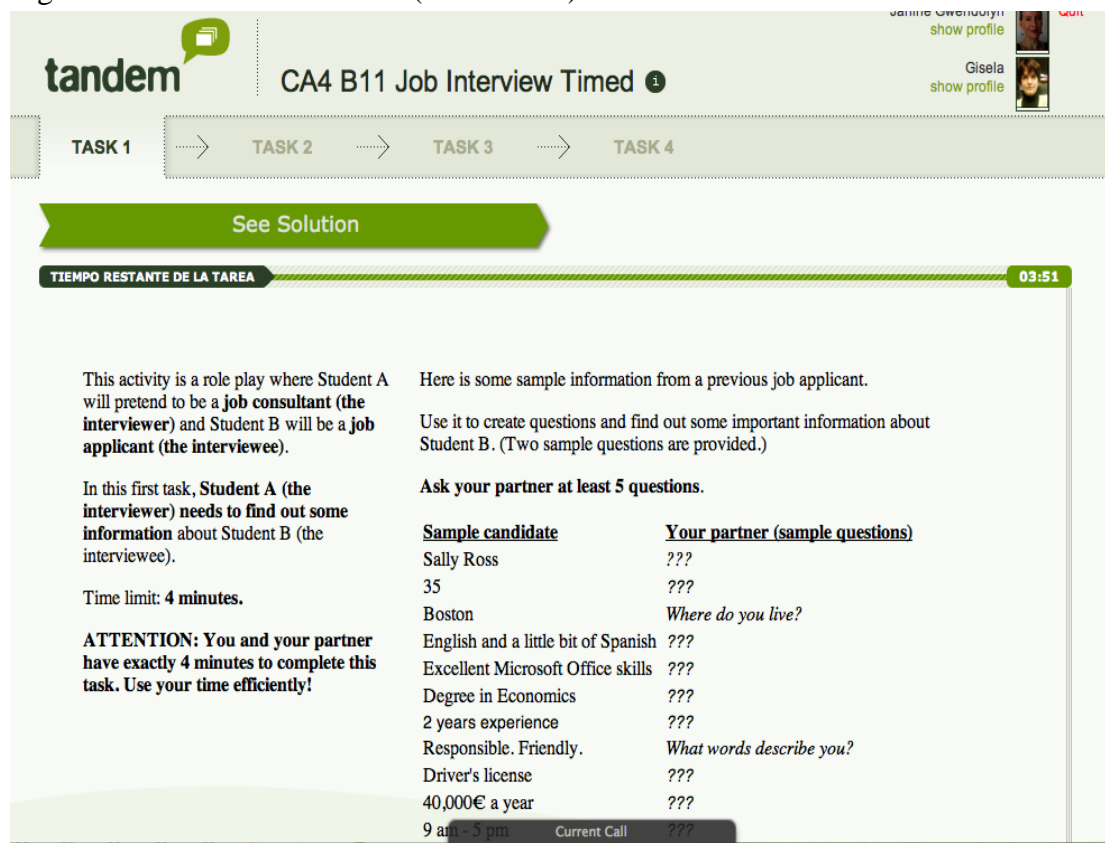
Example 3 (case 3): question and answer

- SILENCE
- 1 L (after 18 seconds) Mhu
  - 2 A (Unintelligible) (PAUSE). I'm waiting. (PAUSE)
  - 3 L Well
  - 4 A Okay, (PAUSE) A or two?
  - 5 L (PAUSE) I have to answer your questions.

- 6 A Okay, perfect.
- 7 L So I'm student B and you are student A
- 8 A Okay, perfect. I start. Well, thank you to be here, to came to this job interview; but I need some information about you. First, what is your name?
- 9 L My name is Laia, Laia Sierra.
- 10 A Okay, How old are you?
- 11 L I'm a thirty-five, sorry thirty-seven years old.

Although learners began their talk somewhat differently they all orientated to each other as interviewer and interviewee with the question and answer turns dominating the talk. None of the participants negotiated their roles as to whether they should be A or B. We propose that this is because learners had been previously assigned roles as interviewer (student A) and interviewee (student B), which had been communicated to them textually through the screen (see figure 3). This assignment of roles was designed to shape which learners carried out verbal turns of questions/answers as interviewer/interviewee as first or second position (talker) in the turn-taking.

Figure 3. Screen for Student A (interviewer)



In addition, based on the analysis of recordings, alongside the screenshots, there also appeared to be instances where learners encountered, orientated to and used the resources in different ways, which we will subsequently demonstrate. This also shaped the talk in different ways. The resources became relevant in their talk in that text from the screen became embedded in oral turns. The text and the symbol '?' were used to support the formation of questions as turns. Resources also became a topic of talk.

a) *Embedded in learners' verbal turns*

One way that learners orientated towards some resources was by embedding them as part of their talk. Example 4 is an excerpt from case 3 and is at the beginning of the recording. Student B cannot hear student A until they have clicked on ‘accept’ to the technological invite initiated technologically by student A.

Example 4 (case 3):

- SILENCE
- 1 L (after 18 seconds) Mhu
  - 2 A (Unintelligible) (PAUSE). **I’m waiting.** (PAUSE)
  - 3 L Well
  - 4 A Okay, (PAUSE) A or two?
  - 5 L (PAUSE) I have to answer your questions.
  - 6 A Okay, perfect.
  - 7 L So I’m **student B** and you are **student A**

After 18 seconds of waiting, in line 2, the learner (interviewer as student A) who must have invited learner (interviewee as student B) using the technological tool *Tandem*, used the phrase ‘I’m waiting’. The turn of his peer is not a response to the previous turn of ‘I’m waiting’ because she could not have heard that turn unless she had accepted his invite before he uttered it (as with the tool in Liddicoat’s study, 2010). In order for them to actually connect, the interaction must be both “technologically and interactionally accomplished” (Tudini, 2014, p. 3). The invite must be carried out somatically by pressing the button. The verbal turn ‘I’m waiting’ which follows 18 seconds later, indicates that while a physical response through the technology has been achieved (because the partner mumbles “Mhu”, line 1), the partner has not yet made an oral turn towards her partner. The word ‘waiting’ in line 2 may have originated from the pop-up on the screen as seen in Figure 4 that also reads ‘waiting’ but we cannot confirm this. We hypothesise that the learner may have said ‘I’m waiting’ in order to update the teacher who will evaluate their recorded interaction or to inform his partner that he is ready to start interacting. The partner’s mumbling may be a form of self-talk because she is orientating herself towards the screen (lines 3 and 5). Raudaskoski (1999) also found that the moment when a learner encountered text was “muttering the words to himself” (Raudaskoski, 1999, p. 115).

Figure 4. Screenshot at start of task

Student A (the job consultant) has two possible jobs available. S/he should explain the jobs to Student B (the candidate).

Student B should ask three questions about the two jobs and then say which job s/he prefers and why.

Time limit: 7 minutes

Here are two different jobs you have available. Briefly describe both of them to your partner. Then answer your partner's questions about the jobs. (If you don't know the answers, be creative!)

Job 1: Bilingual Sales Manager	Job 2: Travel agent
London, England, New York, NY, or Madrid, Spain	Any major city in Europe
Must be willing to travel to other cities once a month	Must be willing to travel to other cities once a week
9 a.m. - 5 p.m. or 12 - 8 p.m.	9 a.m. - 6 p.m.
Must have a car	No car required
£30,000 a year + commission	30,000€ a year

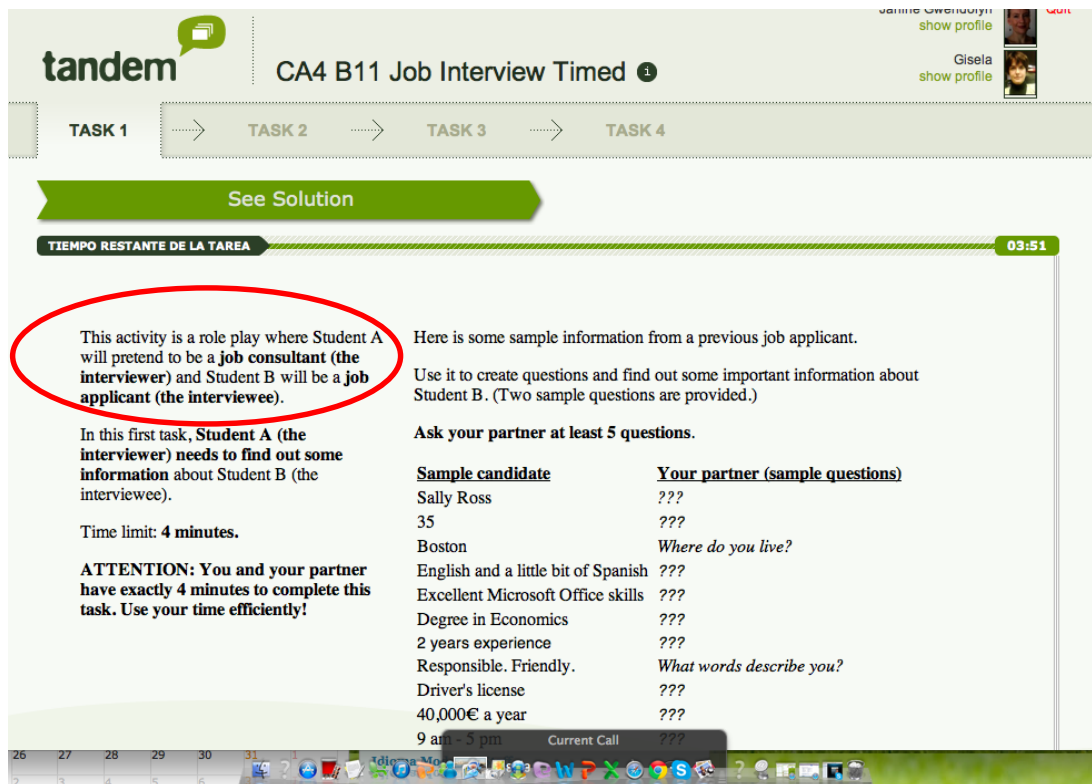
This is a timer based task, please confirm to start:  
It will begin when both you and your partner confirm by clicking the "Start" button.

Waiting...

This result also revealed two things: the act of one learner inviting his partner using the 'start' button showed how learners can carry out turns somatically, echoing Liddicoat's (2010) finding that technology can be a co-constituent of the interaction. The second is the possibility that the screen-based text (the word 'waiting') was being read while waiting for his peer's turn, being re-semiotized by the learner orally. We return to this potential phenomenon further on.

The textual information about the assignment of roles also appeared to embed itself in one of the learner's turns when the invited learner helped to orientate her partner by confirming: 'So I'm student B and you are student A' as shown in figure 5, line 7. This textual information is used by the learner to help orientate her partner towards their role. Screen-text becoming embedded in oral turns also occurs with other screen-based resources which we subsequently present. We understand this use of resource (embedding text into oral turns) as learners orientating to 'others as signs'.

Figure 5. Textual information about students' roles



*b) resources to initiate and form oral turns*

Many learners orientated towards other textual resources (interviewers in tasks 1 and 2) as a support for question formation. Learners (interviewers) asked their partners questions with many learners following a common order from the screen e.g. name, age, where they lived. Checking with the screenshot, the questions frequently asked by cases and the order they are asked, generally corresponded with the screen-based textual resources. Specifically, learners responded to the combination of resources which were: the textual request to act ('Ask your partner at least 5 questions'), the eleven textual resources ('sample candidate' and 'sample questions') combined with the sign '?'. Learners did not deviate or create questions very different from the text sample given, although some interviewers changed the question order or omitted questions in their turn-taking. Whilst the textual request to act can be understood as 'other as discourse' (speech act), the other textual resources for asking questions provide a supporting role for learners in the forming of oral turns.

Examples 5 and 6 show how these resources were used by the interviewer in Task 1 to support help him/her form questions. The bold and underlined text in the transcript shows how the questions correspond to the resources (the sign '?' and the sample answer) in the (edited) screenshot (figure 6).

<p>Example 5 (case 1):</p> <p>1 M Hi</p> <p>2 O How are you today?</p> <p>3 M I'm fine thanks</p>	<p>Figure 6. Resources for asking questions</p>
---	---



<p>4 O Erm Sorry, <b>what is your name?</b></p> <p>5 M My name is Maura, Maura Bater</p> <p>6 O Okay...</p> <p>7 M Yeah</p> <p>8 O sorry... <b>how old are you?</b></p> <p>9 M I'm 34 years old</p> <p>10 O Er Well, <b>where do you live?</b></p> <p>11 M I'm live...I live.. I'm living near <i>of</i> Girona.</p>	<p><b>Ask your partner at least 5 questions.</b></p> <table border="1"> <thead> <tr> <th><u>Sample candidate</u></th> <th><u>Your partner (sample questions)</u></th> </tr> </thead> <tbody> <tr> <td>Sally Ross</td> <td>???</td> </tr> <tr> <td>35</td> <td>???</td> </tr> <tr> <td>Boston</td> <td><i>Where do you live?</i></td> </tr> <tr> <td>English and a little bit of Spanish</td> <td>???</td> </tr> <tr> <td>Excellent Microsoft Office skills</td> <td>???</td> </tr> <tr> <td>Degree in Economics</td> <td>???</td> </tr> <tr> <td>2 years experience</td> <td>???</td> </tr> <tr> <td>Responsible. Friendly.</td> <td><i>What words describe you?</i></td> </tr> <tr> <td>Driver's license</td> <td>???</td> </tr> <tr> <td>40,000€ a year</td> <td>???</td> </tr> <tr> <td>9 am - 5 pm</td> <td>Current Call</td> </tr> <tr> <td></td> <td>???</td> </tr> </tbody> </table>	<u>Sample candidate</u>	<u>Your partner (sample questions)</u>	Sally Ross	???	35	???	Boston	<i>Where do you live?</i>	English and a little bit of Spanish	???	Excellent Microsoft Office skills	???	Degree in Economics	???	2 years experience	???	Responsible. Friendly.	<i>What words describe you?</i>	Driver's license	???	40,000€ a year	???	9 am - 5 pm	Current Call		???
<u>Sample candidate</u>	<u>Your partner (sample questions)</u>																										
Sally Ross	???																										
35	???																										
Boston	<i>Where do you live?</i>																										
English and a little bit of Spanish	???																										
Excellent Microsoft Office skills	???																										
Degree in Economics	???																										
2 years experience	???																										
Responsible. Friendly.	<i>What words describe you?</i>																										
Driver's license	???																										
40,000€ a year	???																										
9 am - 5 pm	Current Call																										
	???																										
<p>Example 6 (case 3):</p> <p>1 A Okay, perfect. I start. Well, thank you to be here, to came to this job interview; but I need some information about you. First, <b>what is your name?</b></p> <p>2 L My name is Laia, Laia Serra.</p> <p>3 A Okay, <b>How old are you?</b></p> <p>4 L I'm a thirty-five, sorry thirty-seven years old.</p> <p>5 A <b>Where do you life?</b></p> <p>Okay, (laughter) me too, no (laughter)</p> <p>6 L I live in Bellver de Cerdanya. In a little town in the Pyrenees.</p>																											

All the participants in the cases used the question ‘Where do you live?’ as shown in figure 6. It can be argued that this textual resource becomes embedded in the talk. However, not all cases used the second complete sentence given: ‘What words describe you?’ highlighting that learners’ choices over how they chose to orientate to textual scaffolds may change from moment to moment. The questions ‘what is your name?’ and ‘how old are you?’ are uttered by the interviewers in most cases. However, these questions do not exist as text on the screen. The symbol ‘?’ and suggested answer ‘35’ (as shown in figure 6 above) provide support for the learner’s turn, which we understand as ‘others as signs’. Learners also modify the text so that for example, ‘city in Europe’ in the text (figure 7) becomes “European city” in example 7, case 2, line 1.

Similar to task 1, in task 2 there are differences in the ways in which the learners orientate towards the screen-based text. In example 7, the interviewer in case 2 uses the two jobs to help structure her talk (lines 1, 3, 5 and 7). She uses the textual information (highlighted in bold in figure 7) to describe the jobs to her partner and also to help her respond to her partner’s questions about the job. Rather than the text initiating or supporting oral turns, the text from the job descriptions become embedded in her talk as we saw previously with the word ‘waiting’ on the navigational resource.

Example 7 (case 2):

- 1G Well err there is enough for me. Err I can offer you two works in our company. Er, the first is a **lingual sales manager** and did require long work experience in that type of sales. Also you said you work two years only. I thing you could be a good candidate. The (pause) second job (pause) is in like a **travel agency**. Um the job is a office in a **European city**, but in this moment I don't know where is the vacant of the travel agency. Um speaking **English is required** but it is also important to speak another **language** fluently like **Spanish** or **French**. The **experience** is a little few; **two years** is good and I don't know what more explain you. Do you have any questions about these jobs?
- 2N Yes... what is the time table?
- 3G Ok. In the **lingual sales managers** is two times to choose: in the morning you work in the **nine am, until five pm**, and yes in the afternoon you can go from **twelve o'clock until eight pm** in the afternoon.
- 4N Ok and what is the salary?
- 5G Is very similar for both jobs: in the **lingual sales manager** is thirty pounds **thirty thousand pounds** sorry a year but there are **commissions**, if you sales (sell) more you earn more, of course. And the travel agent there are **thirty thousands a year** too, but is a salary fixed no by commission.
- 6N Ok, and about first job, I should to travel every day or every week?
- 7G No, only you must be **travel once a month and go to cities... big cities like London, New York and Madrid once a month**, because **English and Spanish** is very important speak fluently.

Figure 7. Textual information for describing two jobs

Here are two different jobs you have available. Briefly **describe both of them** to your partner. Then **answer your partner's questions about the jobs**. (If you don't know the answers, be creative!)

Job 1: Bilingual Sales Manager	Job 2: Travel agent
London, England, New York, NY, or Madrid, Spain	Any major city in Europe
Must be willing to travel to other cities once a month	Must be willing to travel to other cities once a month
High school diploma required	High school diploma and business school required
Spanish and English required	English required; a foreign language (Spanish, French or Chinese) is highly desired
10+ years of sales experience required	2 years experience required in the travel industry or in business
Microsoft Word, Excel, Powerpoint and Access required	Microsoft Word, Excel, and Powerpoint required
Must be a good salesperson	Must be a good communicator
9 a.m. - 5 p.m. or 12 - 8 p.m.	9 a.m. - 6 p.m.
Must have a car	No car required
£30,000 a year + commission	30,000€ a year

In contrast, in example 8, we can see how the interviewer in case 4, does not use the textual job descriptions given.

Example 8 (case 4):

- 1 T And what is your professional experience (unintelligible)

- 2 M Errrr eight years experience with students six to twelve years old
- 3 T What are your personal qualities?
- 4 M I'm very responsibly and (PAUSE) I think that I'm friendly and I have a lot of patience and I'm a sociable person.
- 5 T How much do you hope to earn?
- 6 M I can work for less one thousand euros.
- 7 T (PAUSE) thank you Mona, I can offer you two different jobs in my school. For one hand I have a place in primary with students six to twelve years old, for the other hand I have a place in secondary with students twelve to sixteen years old (yes, Mona interrupts) in two years you will be in this teacher (Mona interrupts) what do you prefer?
- 8 M Mmmmmmm Montse I have a question before I decide.... Mmmm which work time table?
- 9 T (PAUSE) well... in primary school you will work at night o'clock to half past twelve. On the secondary school you'll work at eight o'clock to five o'clock PM
- 10 M Mmm and ... how many groups of students will I have?
- 11 T (PAUSE) in primary school you will have three groups... of students... (cough) Sorry... and in secondary school eight per group.
- 12 M mmm and one question... how much will I earn?
- 13 T (PAUSE) you will earn eight hundred euros in primary school and in secondary school your salary will be two thousand euros (Mona makes humming noises)
- 14 M Mmm for me two jobs are very interesting but I prefer to work... to be an English teacher in secondary school because I haven't any experience high school and I would like to work in secondary school...
- 15 T (PAUSE) Oh very good, Mona. Is fantastic.  
The job is for you! I hope you next Monday at 9 o' clock in my school

Following task design, at this point of the task the screen-based text (job descriptions pertaining to a bi-lingual sales manager or travel agent) should be the same as case 2, in example 7. However, what becomes apparent in case 4's talk is that they did not use the job descriptions provided. This omission shaped the interaction. Instead, the interviewer asserted (previous to the extract presented) that "I'm a headmistress in a private school in Barcelona" and asked questions related to this job description. However, there is some evidence that the interviewer may have used the textual resources from the previous screen (figure 6) to form questions because she refers to "experience" (line 2) and "how much do you hope to earn?" (line 5).

On the one hand, the choice not to use the screen-based job descriptions afforded the learners some creativity, as they created their own jobs. On the other hand, what starts to emerge in the turns and turn-taking is that the interviewee begins to ask questions to the interviewer. The roles of who has 'first position' as interviewer in the turn-taking process and who is the responder (roles of A and B) become blurred. The interviewee orientates to her partner as the questioner (line 8), contravening her pre-given role. This result suggests that the learners' roles of A and B (interviewer and interviewee presented to learners on a previous screen), were not being adhered to. The interviewee 'took the floor' (beyond her role) assuming the role of questioner rather than responder (lines 10 and 12). Therefore, screen-based resources (the communication of roles: who is A and B) may act as a regulatory and/or guiding presence in the absence of a teacher during tasks. Learners' choices to modify their given roles (communicated through different screens) may support possibilities for greater creativity between learners but may also allow for role reversal or allow for oral turns (and time talking) to be less equitable.

## But the computer say me the time is up

The results echo Raudaskoski's (1999) finding that the material/visual semiotics can become "embedded in the sequential meaning making" (Raudaskoski, 1999, p. 132). They appear to initiate learners to orally act (task instruction) and also are orientated towards as support for turn formation. Learners can orientate to some textual/visual resources in a way that the making of verbal turns and meaning making becomes co-shared between the text and the talker, which then enters the discourse. Learners can also choose not to use some textual/visual resources in their negotiation of verbal turns and this may have both positive and negative consequences for task-based talk.

### c) Form part of the topic of talk

On occasions the resources become a topic of the talk between learners. In example 9 below, the dyads have come to the end of Task 1 and the interviewer has asked the questions to the interviewee. The excerpt starts with a response from the interviewee to how much she hopes to earn (line 1).

#### Example 9 (case 3):

- 1 L Oh, well...I have not great aspiration and perhaps a thirty thousand... Euros per year.  
2 A Okay, perfect! Okay, thank you. I look for the position we have. Okay... well ok..now... **close**, no, I suppose...?  
3 L I suppose too, I **close** it...  
4 A Well here are the **questions** and the **answers**... We go to the **next task**...?  
5 L Sorry?  
6 A We go to the **next task**?  
7 L Okay, huh...But the computer say me that the **time is up**.  
8 A Ah, yes.  
9 L So, I think, I think, Tandem is not working correctly.  
10 A No, maybe no because I'm in another page, is in the **solutions page**.

Figure 8. Time Up! Pop-up

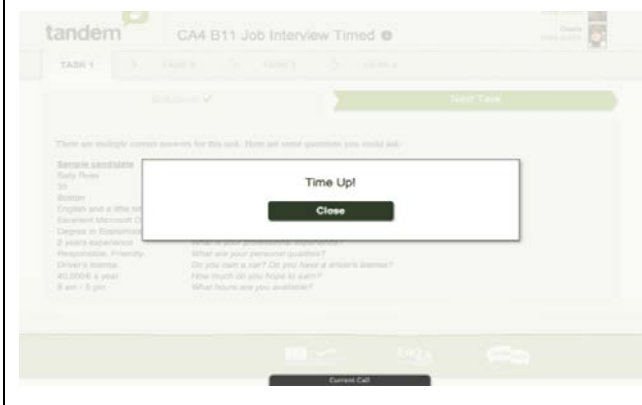
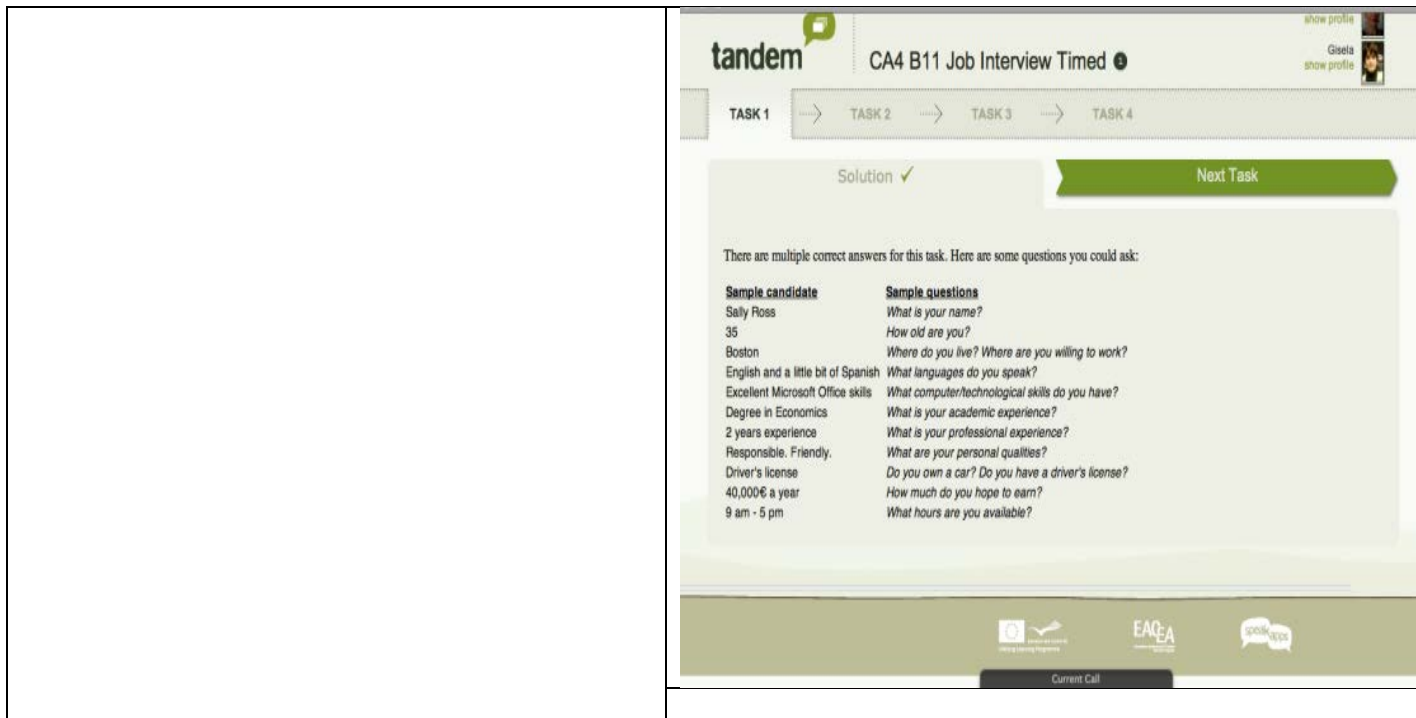


Figure 9. Solutions page



As the learners neared the end of the pedagogical task, learner A says in line 2, “Okay... well ok..now... **close**, no, I suppose...?” to which his partner responded in line 3, “I suppose too, I **close** it...” Because the mention of closing is a new topic of talk, checking the screenshot gives insight into what the meaning may be. It becomes evident that the learners were referring to the Time Up! pop-up which has the text ‘close’ on it as shown in figure 8 above. The word ‘close’ became a topic of talk as both learners engage in a brief discussion as to whether they should close the ‘Time Up!’ pop-up or not. While the word ‘close’ can be interpreted as an ‘invite’ or ‘other as discourse/speech act’ the learner orientated to it as an instruction to act that needs to be negotiated orally with their partner with respect to how they might respond (“**close**, no, I suppose...?” in line 2). They would need to click individually (physically) but negotiate jointly (orally) so that both learners click simultaneously in order to be on the same page. This is possibly made a more difficult task by the fact that they cannot see each other’s gestures or gaze.

Both learners used the word ‘close’ orally (line 2 and 3) as a verb and one learner used the verb ‘go’ with ‘Next Task’ (lines 4 and 6). This suggests that learners orientated to the resources as something that they needed to deal with physically, echoing the *procedural* message type (furthering task completion) identified by Kraut, Fussell and Siegel (2003). The ‘Next Task’ button invited them both to respond physically and they negotiated orally as to whether they are going to carry out that physical (or somatic) response as a turn. The navigational resources become ‘third others’, made relevant in the talk by the learners when they needed to act simultaneously. The need for learners to multi-task during activities (Knight and Barberà, 2016) with respect to carrying out navigational moves, while attending to the pedagogical tasks becomes evident.

This is followed by the partner’s reference to the Time Up! pop up screen resource as an ‘it’ (line 3). The learner is explicit referring to ‘it’ having orientated to the resource as ‘a third other’ that is requesting him to ‘close’. By clicking on ‘close’ we understand that he would be accepting that invite technologically. This echoes Raudaskoski’s (1999) finding of how learners

referred to the computer/program using the term 'this one'. It is "ambiguous in human/inhuman respect, but which nevertheless objectifies the 'other' into a non-copresent entity" (Raudaskoski, 1999, p. 21). The resource is orientated to as 'other as social entity'. The reference to 'it' (line 3) also reflects Kraut, Fussell and Siegel's (2003) *referential* message type (identification/location of task objects).

Following on from the oral negotiation of a navigational act, student A in line 4, indicated that he was on the page with questions and answers (which we hypothesise to be the solutions page, figure 9). He asked his partner whether they should go to the next task. The 'Next Task' button is also on the solutions page and is a textual/visual/navigational resource. However, it becomes clear with his partner's "Sorry?" that she was not saying sorry because she did not hear him, but rather because she did not understand what he was talking about. She says that "The computer say me the time is up" (line 7), indicating that she was not seeing the same resources (therefore nor the same interface) because she was on "another page" (line 10). Ultimately, she decided that the tool was not working correctly. The learner's comment "Sorry?" (line 5) also pertains to Kraut, Fussell and Siegel's (2003) *internal state* message type (relating to intentions, knowledge, emotions etc.).

Apart from the resources ('Close' and 'Next Task') that learners must use to navigate, both learners also made reference to the resources in order to communicate and check what page they were looking at on the screen. The use of deixis ("here") emphasised location and a need to highlight learners' proximity in relation to a screen-based resource, another *referential* message from Kraut, Fussell and Siegel's message types (2003). "Well here are the questions and answers" (line 4) and later his partner said "I'm in another page, is in the solutions page" (line 10). The words "here" (line 4) and "I'm in" (line 10) emphasise learners' apparent need to establish and check what their partner is seeing. Their talk becomes a tool for physical/visual orientation for them. Their use of verbal summary of what is on the screen ("here are the questions and answers" and "in the solutions page") rather than direct referencing of the resources on the screen, appears to suggest that once learners realize they are not on the same page or have 'lost' each other, the mediation of smaller screen-based resources, which were important in the pedagogical task, are no longer useful. Their lens for mediating to problem solve has expanded outwards to include mediation through verbal references with the interface 'pages' as a means to share mutual understanding about their problem. We liken this to the expanding aperture of a camera lens. The multimodal ensemble is not so important at this moment but rather the interface page itself becomes the important resource.

Thus, in relation to research question 2, regarding how learners initiate or respond to other screen-based resources (and what this tells us about turn-taking), we identified that learners respond to other screen-based resources orally and somatically which is made relevant in their oral (and physical) turns. Screen-based resources also appear to initiate turns (task instruction to 'ask questions') and navigational resources can become a topic of talk. In addition, learners must 'initiate' with some resources physically ('start' and 'next task'). These can be considered passive agents because they respond to a learner's initiation (to move learners to another place) through touch (the somatic mode). Learners may also need to 'accept' as a response to navigational resources such as pop-ups.

We now turn to the pop-up resource which we identify as an 'active agent' (Dourish, Bellotti, Mackay and Ma, 1993) and turn-taker in the discourse, adding to our understanding of how learners respond to pop-ups as 'others as (active) agents'.

### **The pop-up: the interrupter and the messenger**

Not only did the text from navigational resources appear to become embedded in learner talk ('waiting', example 4, case 3) or become a topic of talk ('close' on the pop-up and 'next task' button in example 9, case 3) but learners orientated to them in other ways. The first is with regard to the Time-up pop-up daemon (figure 8), which is orientated towards by learners as an interrupter of a turn and secondly, as a deliverer of a message. Both ways highlight that learners can orientate towards navigational resources as 'others as social entities' or agents.

#### *The interrupter*

In example 10 below, the learners were coming to the end of Task 2 and the interviewee is clarifying some aspects of the two jobs that her partner, the interviewer, had presented her with. After declaring "Yes, I prefer the second job" (line 7), she follows with "Time Up! Ok" (line 7) and continues in the same turn with why she prefers the second job. The learner orientated towards the pop-up as a brief interrupter of the conversation with her partner, which she noted/acknowledged ("ok" in line 7) and which the learners manage, calmly, in order to continue the closure of their interview.

Example 10 (case 1):

- 1 L You have said I need a diploma or qualification, concrete qualification, *specially* (special) qualification? Really?
- 2 S Yes, diploma, high school diploma. And business school
- 3 L Business school... okay, okay. Well without commission... and the commissions... do you *have* about what commissions can be?
- 4 S I'm sorry I don't, I don't know.
- 5 L Okay, well...
- 6 S Do you prefer?
- 7 L Yes, I think I prefer the second job. **Time up!** Ok, well I prefer the second job but I don't know if I can choice that because I don't have the experiences in industries but for me it's better, okay?
- 8 S Okay good, thank you very much for your time, it's a pleasure er meeting you today (interference).
- 9 L Okay me too thanks

#### *The messenger*

In another case, a learner orientated to the same resource differently, as shown in example 9, case 3 previously. The learner said "Okay, huh...But the computer say me that the time is up". The 'other' here (this time referred to as "the computer") is referred to like a messenger. The 'message' only appeared on her screen (following task design) which she then conveyed to her partner as an explanation as to why she could not fulfil her partner's request that they move to the next task.

### **7. Conclusions**

This study has explored how the online nature of the participants' interaction was a relevant constituent part of learners' interactions (following Liddicoat, 2010), not just a facilitative channel for interaction. The focus on semiotic resource-as-mode(s) rather than mode-as-channel has played an important role in highlighting how learners' various orientations to screen-based resources can lead to differences in turns and turn-taking. The resources shaped learners' talk by becoming embedded, modified in oral turns; initiating and supporting oral turns; becoming topics of talk as well as orientated towards as turn-takers. Turn-takers can be 'active agents' (Dourish, Bellotti, Mackay and Ma, 1993) or passive agents. Learners are 'semiotic initiators and/or responders' (Coffin and Donohue, 2014) towards 'others' as text, social entity/agent and signs through oral and/or somatic turn-taking. Some resources are also found to 'initiate'.

Future research on turn-taking therefore, might consider not only how learner generated turns can be produced 'transmodally' through speech and text (Lamy, 2006) but that screen-based resources can be potential 'others', acting as and/or be oriented towards by learners as discourse participants (as well as resources), contributing to a multi-party accomplishment of the discourse. Learners may need to initiate/respond to screen-based turns in some way, all be it that some resources may be 'sporadically conversational' (Herring, 2005).

This study highlights the possibility of recognizing and analysing a scenario that offers learners 'a multimodal experience' (Collentine, 2009), moving the focus beyond the oral aspects of talk in the co-construction of meaning that are predominantly considered in SLA research.

Attending to the 'material stuff' (Kress, 2003) of the screen has been important for understanding how oral turns and meaning are/is made. The existence and relevance of the combination of embodied (voice) and non-embodied modes and resources are highlighted. Future studies on spoken tasks whereby physical activity is occurring simultaneously to learner talk might look to frameworks in the HCI literature such as Kraut, Fussell and Siegel's (2003) when analysing spoken interaction, particularly the process of meaning negotiation.

## **Bibliography**

Austin, J.L. (1962). *How to Do Things with Words*. Cambridge: Harvard University Press.

Benson, P. (2015). You Tube as text: spoken interaction analysis and digital discourse. In Jones, R. H., Chik, A., & Hafner, C. A. (Eds.), *Discourse and digital practices: doing discourse analysis in the digital age* (pp. 81- 96). Oxford and New York: Routledge.

Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: a social semiotic frame*. London: Routledge.



- Blake, R. (2016). Technology and the four skills. *Language Learning & Technology*, 20(2), 129-142.
- Block, D. (2013). Moving beyond Lingualism: Multilingual Embodiment and Multimodality in SLA. In May, S. (Ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 54-77). New York: New York.
- Calvo-Ferrer, J. R., Melchor-Couto, S., & Jauregi, K. (2016). ReCall Special Issue: Multimodal Environments in CALL Editorial Multimodality in CALL. *ReCALL*, 28(3), 247-252.
- Clark, H.H., & Brennan, S.E. (1991). Grounding in Communication. In L.B. Resnick, J. H. Levine, and S.D. Teasley. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, (pp. 127-149).
- Coffin, C., & Donohue, J. (2014). A language as social semiotic-based approach to teaching and learning in Higher Education, *Language Learning*, 64(1), 1-212.
- Collentine, K. (2009). Learner use of holistic language units in multimodal, task-based synchronous computer-mediated communication. *Language Learning & Technology*, 13(2), 68-87.
- Coulthard, M. (1985) *An Introduction to Discourse Analysis*, 2<sup>nd</sup> edn (1<sup>st</sup> edn 1977), London: Longman.
- Dourish, P., Bellotti, V., Mackay, W., & Ma, C. Y. (1993, December). Information and context: lessons from the study of two shared information systems. In *Proceedings of the conference on Organizational computing systems* (pp. 42-51). ACM.
- Gilbert, R., Manchón, R. & Vasylets, O. (2016). Mode in Theoretical and Empirical TBLT Research: Advancing Research Agendas. *Annual Review of Applied Linguistics*, 36, 117-135.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (pp. 41-58). New York: Academic Press.
- Guth, S., & Helm, F. (2011). Developing multiliteracies in ELT through telecollaboration. *ELT journal*, 66(1), 42-51.
- Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(02), 116-137.
- Jenks, C.J. (2014). *Social Interaction in Second Language Chat Rooms*. Edinburgh: Edinburgh University Press.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing Multimodality*. London: Routledge.
- Knight, J. & Barberà, E. (2016). The Negotiation of Shared and Personal Meaning making in Spoken Interaction Tasks. In Pixel (ed.). *Conference Proceedings of the 9th International*

Conference of ICT and Language learning (pp. 248-252). Italy: Librariauniversitaria.it edizioni.

Authors (2017 forthcoming). Navigating a multimodal ensemble: learners mediating verbally and non-verbally in online interaction.

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-Computer Interaction*, 18, 13-49.

Kress, G. (2003). *Literacy in the new media age*. London: Routledge.

Lamy, M.-N. (2006). Multimodality in second language conversations online: Looking for a methodology. In proceedings of third international conference on multimodality (pp. 385- 403). Pavia, Italy.

Levy, M., & Kennedy, C. (2004). A task-cycling pedagogy using audio-conferencing and stimulated reflection for foreign language learning, *Language Learning and Technology*, 8(2), 50-68.

Liddicoat, A. J. (2010). Enacting participation: Hybrid modalities in on-line video conversation. In C. Develotte, R. Kern, & M. N. Lamy (Eds.), *Décrire la conversation en ligne* (pp. 37–50). Lyon, France: ENS Editions

Liou, H. (2011). Blogging, Collaborative writing, and Multimodal Literacy in an EFL context. In M. Levy, F. Blin, C. B. Siskin, & O. Takeuchi (Eds.), *WorldCALL: International perspectives on computer-assisted language learning* (pp. 3-18). London: Routledge.

Norris, S. (2006). Multiparty interaction: a multimodal perspective on relevance. *Discourse studies*, 8(3), 401-421.

Raudaskoski, P. (1999). *The use of communicative resources in language technology environments. A conversation analytic approach to semiosis at computer media* Doctoral dissertation. Retrieved from <http://vbn.aau.dk/ws/files/72280035/pirkkosphd.pdf>

Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation, *Language*, 50(4):696-735.

Searle, J. (1969). *Speech Acts*, Cambridge: Cambridge University Press.

Sellen, A. J. (1995). Remote conversations: The effects of mediating talk with technology. *Human-computer interaction*, 10(4), 401-444.

Stenström, A. B. (1994). *An Introduction to Spoken Interaction*. London & New York: Longman.

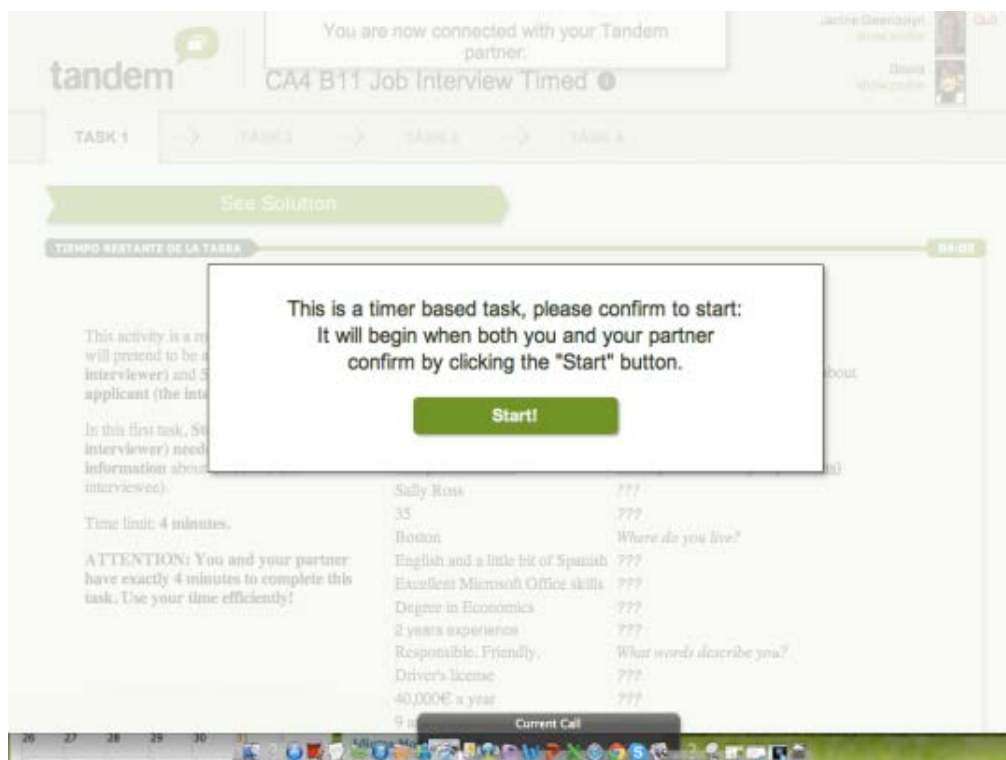
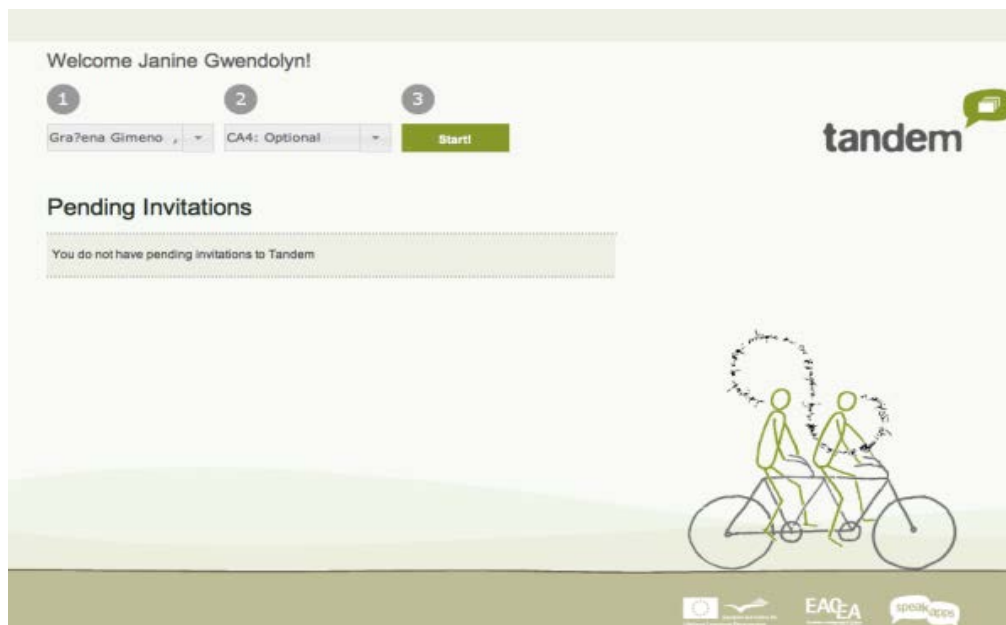
Tudini, V. (2014). Conversation analysis of computer-mediated interactions. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-7). Hoboken, NJ, USA: John Wiley & Sons.

Yamada, M. (2009). The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: Experimental study. *Computers & Education*, 52(4), 820-833.

## **Appendices**

*Appendix A: A sample of screenshots from the reconstructed task collected from researcher simulation*

But the computer say me the time is up



tandem CA4 B11 Job Interview Timed 1

TASK 1 TASK 2 TASK 3 TASK 4

See Solution

TIEMPO RESTANTE DE LA TAREA 03:51

This activity is a role play where Student A will pretend to be a **job consultant (the interviewer)** and Student B will be a **job applicant (the interviewee)**.

In this first task, **Student A (the interviewer)** needs to find out some **information** about Student B (the interviewee).

Time limit: **4 minutes**.

**ATTENTION: You and your partner have exactly 4 minutes to complete this task. Use your time efficiently!**

Here is some sample information from a previous job applicant.

Use it to create questions and find out some important information about Student B. (Two sample questions are provided.)

**Ask your partner at least 5 questions.**

Sample candidate	Your partner (sample questions)
Sally Ross	???
35	???
Boston	Where do you live?
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible, Friendly.	What words describe you?
Driver's license	???
40,000€ a year	???
9 am - 5 pm	???

3.

tandem CA4 B11 Job Interview Timed 1

TASK 1 TASK 2 TASK 3 TASK 4

See Solution

TIEMPO RESTANTE DE LA TAREA 02:35

This activity is a role play where Student A will pretend to be a **job consultant (the interviewer)** and Student B will be a **job applicant (the interviewee)**.

In this first task, **Student A (the interviewer)** needs to find out some **information** about Student B (the interviewee).

Time limit: **4 minutes**.

**ATTENTION: You and your partner have exactly 4 minutes to complete this task. Use your time efficiently!**

Here is some sample information from a previous job applicant.

Use it to create questions and find out some important information about Student B. (Two sample questions are provided.)

**Ask your partner at least 5 questions.**

Sample candidate	Your partner (sample questions)
Sally Ross	???
35	???
Boston	Where do you live?
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible, Friendly.	What words describe you?
Driver's license	???
40,000€ a year	???
9 am - 5 pm	???

4.

But the computer say me the time is up

tandem CA4 B11 Job Interview Timed show profile

TASK 1 → TASK 2 → TASK 3 → TASK 4

See Solution

TIEMPO RESTANTE DE LA TAREA 01:16

This activity is a role play where Student A will pretend to be a **job consultant (the interviewer)** and Student B will be a **job applicant (the interviewee)**.

In this first task, **Student A (the interviewer)** needs to find out some information about Student B (the interviewee).

Time limit: **4 minutes**.

**ATTENTION: You and your partner have exactly 4 minutes to complete this task. Use your time efficiently!**

Here is some sample information from a previous job applicant.

Use it to create questions and find out some important information about Student B. (Two sample questions are provided.)

Ask your partner at least 5 questions.

Sample candidate	Your partner (sample questions)
Sally Ross	???
35	???
Boston	Where do you live?
English and a little bit of Spanish	???
Excellent Microsoft Office skills	???
Degree in Economics	???
2 years experience	???
Responsible, Friendly,	What words describe you?
Driver's license	???
40,000€ a year	???
9 am - 5 pm	???

Current Call

5.

tandem CA4 B11 Job Interview Timed show profile

TASK 1 → TASK 2 → TASK 3 → TASK 4

Solution ✓ Next Task

There are multiple correct answers for this task. Here are some questions you could ask:

Sample candidate	What are your professional experiences?
Sally Ross	What are your personal qualities?
35	Do you own a car? Do you have a driver's license?
Boston	How much do you hope to earn?
English and a little bit of Spanish	What hours are you available?
Excellent Microsoft Office skills	
Degree in Economics	
2 years experience	
Responsible, Friendly,	
Driver's license	
40,000€ a year	
9 am - 5 pm	

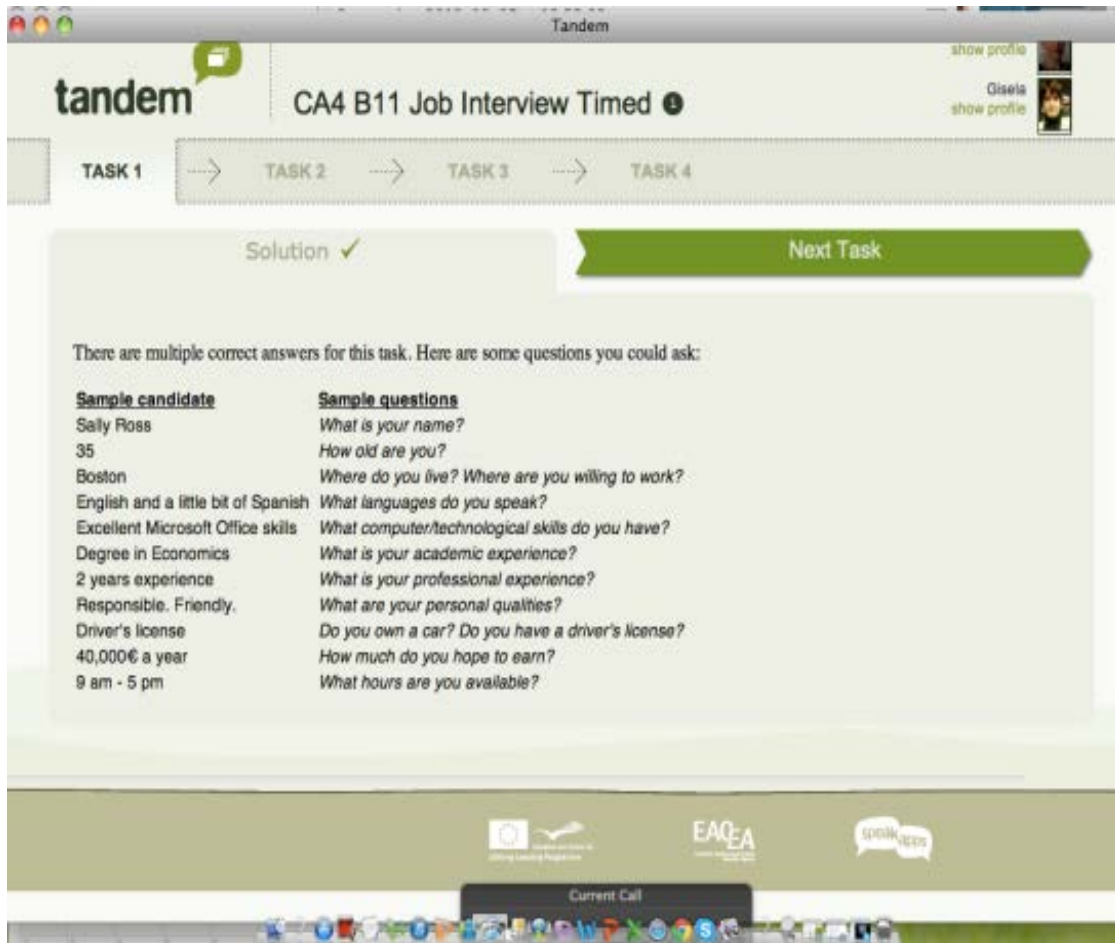
Time Up!

Close

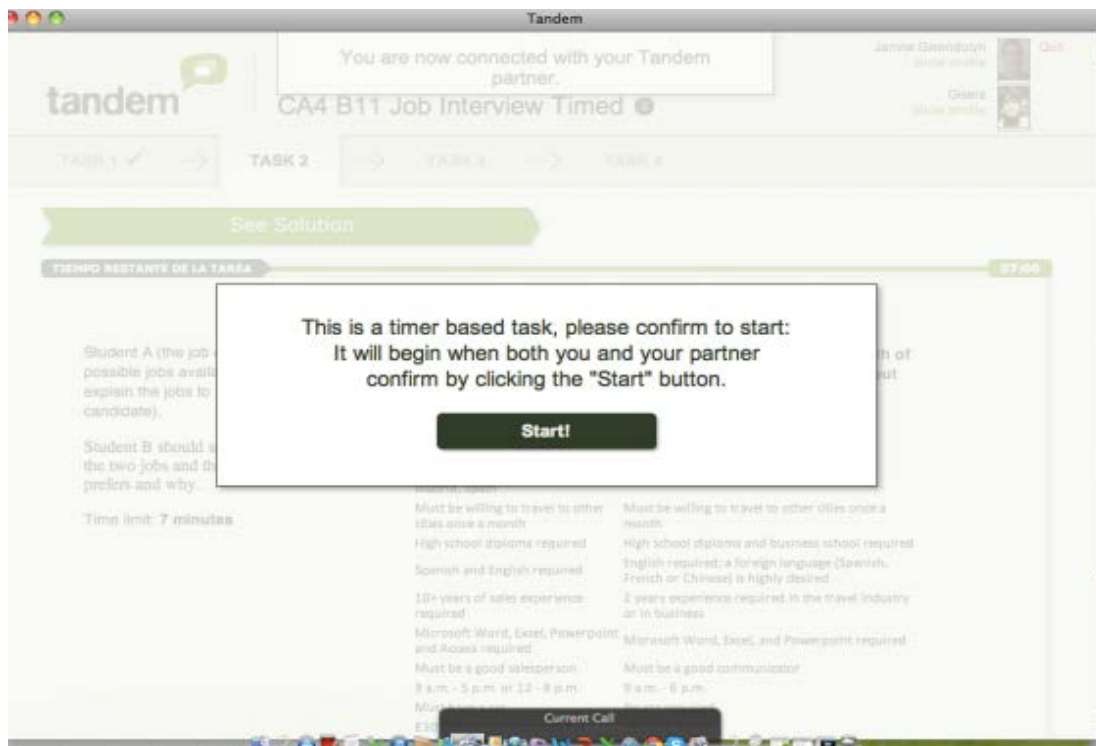
Current Call

6.

7.



8.



The screenshot shows a web browser window titled "Tandem" with a navigation bar containing "TASK 1 ✓", "TASK 2", "TASK 3", and "TASK 4". Below the navigation bar is a green button labeled "See Solution". A progress bar at the top indicates "TIEMPO RESTANTE DE LA TAREA" with a timer set to "09:56".

Student A (the job consultant) has two possible jobs available. S/he should explain the jobs to Student B (the candidate).

Student B should ask three questions about the two jobs and then say which job s/he prefers and why.

**Time limit: 7 minutes**

Here are two different jobs you have available. Briefly **describe both of them** to your partner. Then **answer your partner's questions about the jobs**. (If you don't know the answers, be creative!)

<b>Job 1: Bilingual Sales Manager</b>	<b>Job 2: Travel agent</b>
London, England, New York, NY, or Madrid, Spain	Any major city in Europe
Must be willing to travel to other cities once a month	Must be willing to travel to other cities once a month
High school diploma required	High school diploma and business school required
Spanish and English required	English required; a foreign language (Spanish, French or Chinese) is highly desired
10+ years of sales experience required	2 years experience required in the travel industry or in business
Microsoft Word, Excel, Powerpoint and Access required	Microsoft Word, Excel, and Powerpoint required
Must be a good salesperson	Must be a good communicator
9 a.m. - 5 p.m. or 12 - 8 p.m.	9 a.m. - 6 p.m.
Must have a car	No car required
€30,000 a year + commission	30,000€ a year

At the bottom of the screen, there is a "Current Call" window and a "speak" button.

9.



### **13. Appendix: additional articles in journals**

Paper 3: Learner agency and its effect on spoken interaction time in the target language

# Learner agency and its effect on spoken interaction time in the target language

**Janine Knight**

Universitat Internacional de Catalunya  
 janine@uic.es  
 Universitat Oberta de Catalunya  
 jknight@uoc.edu

**Elena Barberà**

Universitat Oberta de Catalunya  
 ebarbera@uoc.edu

*This paper presents the results of how four dyads in an online task-based synchronous computer-mediated (TB-SCMC) interaction event use their agency to carry out speaking tasks, and how their choices and actions affect time spent interacting in the target language. A case study approach was employed to analyse the language functions and cognitive and social processing that occurred in audio recordings of spoken interaction between four dyads, alongside other indicators of pre-task behaviour, triangulated with results from learner questionnaires. The study revealed that whilst all cases engaged in overt spoken interaction, some cases also avoided the designed task and engaged in covert pre-task planning. Learners' ability to reconfigure 1) the time mode of the task design; 2) the ways in which technological tools were used and 3) language choice, all impacted on their time spent interacting in the target language. The findings highlight tensions between learners' choices across the three dimensions that they had reconfigured, raising questions as to how to support time in synchronous interaction in the target language whilst supporting learners' agency. The implications are presented and discussed.*

**Keywords:** language choice, task avoidance, time mode, task-based synchronous computer-mediated communication (TB-SCMC), spoken interaction, learner agency

## 1. Introduction

Increasing global access to and widespread use of Information Communication Technologies means that new tools to facilitate synchronous spoken interaction are emerging. This in turn implies potentially 167

new and diverse opportunities for learners to exercise agency whilst practicing their speaking skills. Although the actual impact of practising speaking skills via synchronous computer-mediated communication (henceforth voice-based SCMC) has not yet been sufficiently examined, oral SCMC still provides excellent opportunities for students to practise their English with flexibility of time and place (Yang & Chang, 2008). Having enough opportunities and time to practice speaking is not only a concern for teachers and designers (Appel, Robbins, Moré & Mullen, 2012; Hampel & Hauck, 2004; O'Dowd, 2000 & Wang, 2006) but also learners (Hurd, 2007). Students can be concerned with developing fluency, having enough practice and finding opportunities to talk to others (Hurd, 2007). Therefore, maximising opportunities for speaking practice is not only an important goal from these perspectives but it is also deemed that language is best learned and taught through interaction, which in turn contributes to language gains (de la Colina & Garcia Mayo, 2007; Gass & Mackey, 2006). Specifically, rapid real-time interaction, which is necessary to develop fluency, can be facilitated through spontaneous synchronous tasks. Not only can tasks be designed to allow for maximum speaking time for students (Stickler, Batstone, Duensing & Heins, 2005) but advances in Voice over Internet Protocol (VoIP) and related tools such as Skype boost the possibilities to facilitate synchronous peer-to-peer spoken interaction in the target language (henceforth TL). Although 'sufficient practice' can be understood as the number of opportunities to speak on offer during a language learning course, it can also mean length of time interacting.

Despite the pedagogical and technological factors that contribute to the task design to support spontaneous synchronous tasks, learners may make and act on choices that may not run in accordance with this goal. Learners may exercise agency in ways that do not necessarily optimise opportunities for spoken interaction in the TL, and which appear contrary to their own desire to practice as part of their (assumed) goal of learning a language. The problem therefore is that although the exercising of agency may have inherent value (Schwartz & Okita, 2009) based on the belief that learners are agents who "play a defining role in shaping qualities of their learning" (Dewaele, 2009, p. 638), learners may use agency in ways that may not make optimal use of the opportunities to interact orally with others in the TL.

This study attempts to explore how students exercise agency in two online tasks and how this affects time spent interacting in the TL. It describes a task-based SCMC event for spoken interaction and the choices learners make in relation to the 'task-as-workplan' (Breen, 1987) or task (design) expectations for synchronous, spontaneous interaction. Some learners' choices result in a reconfiguration of task dimensions that run contrary to task-as-workplan. These are: 1) language choice of either first language (L1) and/or the TL; 2) choices relating to technological tools and 3) choices relating to time mode (i.e. synchronous and/or asynchronous). Specifically, we are interested in learners' choices and actions in relation to each dimension and how they intersect. Because the interaction time in the TL is important for teachers, designers and learners, we also focus on how these three dimensions affect this task outcome.

## 2. Theoretical Framework

### 2.1 Learner agency, speech and task-based learning

There is a plethora of definitions regarding learner agency. Martin's (2004) definition as "the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives" (p. 135) is used here over the more common "socioculturally mediated capacity to act" (Ahearn, 2001, p. 112). Although many agree that an individual's capacity to act is socioculturally, contextually and interpersonally mediated (Mercer, 2011), Martin's definition allows for a focus on agency as intentional behaviour during task processes. It goes beyond the conceptualisation of agency as a capacity to being a property of the individual. Agency is a fundamental construct in understanding learning processes and learner identities (Miller, 2012), and many perspectives agree that language plays a central role in current thought on agency. At the level of speech, language can be used as an action or process (Swain, 2006) alongside which, learners can use their motor and sensory systems in order to carry out agentic actions (Bandura, 1999). In online tasks, this may mean using language for problem solving or responding to textual instructions, as well as physically navigating through a task for task completion.

Regarding tasks, Breen (1987) highlighted differences between 'task-as-workplan' (concerned with expectations and intentions of task design) and 'task-in-process' (what learners actually do). Conceptualising pedagogical tasks in these two ways helps to highlight "the notion that learners, as active agents in learning processes, can modify activities according to their own intentions – modifications which may or may not be in direct accordance with the initial intentions of that 'task-as-workplan'" (Dooly, 2011, p. 72). Task design intentions and the expectations arising from them can also be conceptualised as 'lines of desire', an architectural term that Lukin and Du Boulay (2003) apply to technology-use to highlight the path or trajectories that people make but which are often shortcuts that ignore the given route.

### 2.2 Language choice and language avoidance as choice

Sociocultural perspectives of language learning have suggested that students learning a new language can use their L1 to serve a variety of functions, which can ultimately support learning processes and outcomes. This can include task management (Macaro, 1997), scaffolding while working on a task (Kötter, 2003), identity negotiation (Myers Scotton, 1983), discussing unknown language words (Knight, 1996) and social interaction (Tarone & Swain, 1995). Language choice and control of language use, including codeswitching between languages, can be understood as a naturally occurring expression of learner agency (Garcia, 2009) forming part of students' multilingual repertoire for different or identical practices. Some research indicates that multilingual practices can contribute to the eventual construction of a final monolingual output (Dooly, 2011) scaffolding cognitive and communicative activities which eventually allow speakers to participate in monolingual activities at the end of the process (Borràs, Canals, Dooly, Moore, & Nussbaum, 2009). Some researchers have found that multilingual language learners working towards monolingual task accomplishment tend to shift between different types (or stages) of L1 and target language use (Borràs et al., 2009; Masats, Nussbaum, & Unamuno, 2007). Their codeswitching allows them to overcome communicative obstacles, facilitating a final stage where the learner can

maximise the use of the TL (Dooly, 2011). A learner's plurilingual repertoire can be understood as "resources for practicing agency, that is, the right to make and to enact their own linguistic choices, in goal-orientated and context-embedded situations" (Vitanova, Miller, Gao & Deters, 2014, p. 8). Through language use multilingual subjects can exercise agency (Kramsch, 2009).

Reasons cited for learners choosing to use the L1 as outlined above may fall within what Schwartz and Okita (2009) call "causal-focused and rights-focused applications of agency" (p. 7). Whereas the former is concerned with what conditions foster learning, the latter is concerned with protecting or enabling people's access to a particular form or expression of learning.

Other research that is concerned with the conditions that foster learning includes studies that conceptualises language choice as an avoidance strategy. For example, in Musk's (2014) research on learners' language choices and use of Google translation tool he found that learners made incremental choices that avoided the TL (or favoured the L1) during computer-assisted project work: learners acted upon their language preference, opting to read in their L1. He also noted that learners tended to be product-oriented, drawing on their previous experience to get the job done quickly and efficiently, including their experience of and familiarity with the technological tools. He proposed that different learners rely more on their need or choice to translate and use translation tools because of differences in confidence and reading proficiency in the TL. Musk (2014) suggests that the study of language choice may have been impacted by Tarone's (1978) study, which groups language switching not as an avoidance strategy but rather as a subcategory of conscious language transfer: the conscious judgement that something in the native language – most typically – and something in the TL are similar, if not actually the same (Odlin, 1989).

We now present other studies that involve factors relating to choice and avoidance. This is in order to extend the literature review to include studies that are also concerned with what conditions foster learning.

### *2.3 Avoidance, spoken interaction and off-screen behaviours*

Learner avoidance of interacting in the TL is commonly studied as a communication strategy during spoken interaction as learners avoid syntactic or lexical items, and topics and concepts that pose language difficulties or pronunciation issues (from Tarone's framework, 1981) including in SCMC oral tasks (e.g. Kim, 2014). Within studies of avoidance as a communication strategy, Horwitz, Horwitz, and Cope (1986) suggested a relationship between unwillingness to interact verbally and Foreign Language Anxiety. Research suggests that this tendency is influenced by factors such as personality traits or overall unwillingness to communicate (Levine, 2003). The current study, however, focuses on avoidance of carrying out the 'task-as-workplan' because this provides us with a more complete understanding of avoidance in online tasks.

Avoidance and tool use were noted by Appel, Robbins, Moré and Mullen (2012), who explored how different interface versions affected spoken interaction tasks designed for spontaneous, synchronous interaction. They found that over 60% of the learners had looked at the materials and prepared beforehand, therefore avoiding the synchronous mode of the task. Students reported being more nervous when working with the interface version (Tandem tool), which required "a degree of improvisation and spontaneity which put

additional pressure on the students" (Appel, Robbins, Moré & Mullen, 2012, p. 18), over another interface version, which allowed preparation.

No other specific studies of avoidance of oral CMC tasks were identified in the research literature. However, there are a number of studies of off-screen and out-of-class behaviours, which are complementary in the sense that they give insight into learner processes. Studies of learners' off-screen behaviours with CMC largely fall into two groups (Suzuki, 2013): online contexts where learners are in dyads with other learners and native speakers of the TL (González-Lloret, 2011; Tudini, 2010), or in their physical environment (Kitade, 2008; Leahy, 2004). In the latter group, Suzuki (2013) studied one learner in front of the computer during a teacher-led synchronous Japanese class mediated by audio-based conferencing software, where class participants' behaviour in their physical environments were not observable to others. Results revealed that the learner gained significant affordances from the online/off-screen course format, allowing her to create her own learning opportunities including actively taking private turns in her physical environment without being heard by others.

Hafner, Li and Miller (2015) studied university learners' non-online out-of-class behaviours in an English course project work. They found that students' computer-mediated interactions (Facebook, WhatsApp and email) were plurilingual, with students drawing on English, Chinese and mixed code to different extents. Different languages were used strategically: whereas L2 was used more in the construction of the final project product, L1 was used more to promote group cohesion.

Hampel and Stickler (2012) also noted off-task conversations between students in a videoconferencing session which they stated usually occurred in the mode other than the one the teacher was using.

Other factors related to language avoidance pertain to learner attitudes and beliefs about L1 use in the classroom. Learners' attitudes can affect the extent and function of own-language use in the classroom and its potential contribution to learning (Hall & Cook, 2012). Beginner university learners were found to prefer L1 use for classroom management and suggested that its use reduces anxiety confirming a positive affective role that L1 use can play (Rolin-Ianziti & Varshney, 2008).

#### *2.4 Interaction time, time mode preference and planning time for tasks*

Interaction has long been acknowledged as one of the most influential factors contributing to language learning (Boonsue, Janssen & Srinaowaratt, 2015). Although many studies on SCMC have focused on the number of turns taken as an indicator of the amount of interaction taking place (Blake, 2005; Jepson, 2005), other researchers have used the amount of time in interaction as a measure of user engagement in the TL (Stickler, Batstone, Duensing & Heins, 2005). Interaction time has been considered by various researchers, (although not using this exact term), in relation to the tension between time spent in the L1 or the TL. Whilst some argue that there is a case for using the L1 in the classroom as learners change languages and L1 use can be very time-efficient in certain situations (Üstünel & Seedhouse, 2005) others argue that speaking in the TL in the classroom should occur as much as possible (Moeller & Roberts, 2013). Levine (2011) proposes that a plurilingual pedagogy, including the positive use of L1 as a language choice, would increase the total interaction time in the TL. He proposes that while the absolute ratio of TL to L1 communication might

decrease, the absolute amount of time spent communicating in the TL would increase because students would talk more.

Apart from the issue regarding learners' L1 use in the language classroom, time interacting in the TL is an important goal in SCMC contexts from learners' (Hurd, 2007), teachers' and designers' perspectives (Appel, Robbins, Moré & Mullen, 2012; Hampel & Hauck, 2004; O'Dowd, 2000; Wang, 2006).

Regarding time mode, SCMC is not flexible (Levy & Stockwell, 2006) since learners engage in 'live' communication with partners, so they must schedule specific times for study. Some researchers consider SCMC or asynchronous (ACMC) time modes as being learner preferences or learning styles (Benbunan-Fich & Hiltz, 2003; Wang, Wang, Wang & Huang, 2006). These are behaviours related to the psychological, cognitive, and affective domains of interaction within learning environments, which also involve learners' preferred ways to receive, process, and recall information during instruction (Aragon, Johnson & Shaik, 2002). Shahabadi and Uplane (2014) found distinct differences in learning styles between learners in different time modes.

The off-screen or private talk (Suzuki, 2013) that learners engage in during online interaction can also be understood in terms of planning time and learner performance in the TL. Types of task planning are differentiated based on when planning occurs namely pre-task planning and online planning (Ellis, 2005). The planning pertinent to this current study is pre-task planning which can be 'strategic planning time' (i.e. deliberation of content and code) or 'rehearsal time' (i.e. a practice run through of the task), both of which can occur prior to task performance (Ortega 1999; Ellis 2005). Planning allows learners to attend to language as form and studying planning gives an insight into what learners attend to, and what effect it has on the way they use language (Ellis, 2005). Although research suggests that planning affects ways in which learners perform a task, there is very little research about what learners actually do when they plan (Ellis, 2005). Batstone (2005) takes a sociocognitive view of planning highlighting that learners can approach tasks in two ways: requiring economical and efficient communication or providing opportunities for engaging in learning activities. Furthermore, with respect to agency and planning, Batstone suggests that some language learners are more 'face sensitive' (Batstone, 2005; Aston, 1986) than others and "exploit principles of clarity/economy to the hilt out of a concern with self-protection rather more than with self-expression" (Batstone, 2005, p. 288). Ellis (2005) suggests that the context of tasks also shapes how learners plan for and perform tasks such as testing conditions.

With regard to pre-task planning, studies have largely demonstrated a benefit for complexity (the use of more advanced or more diverse TL features), accuracy (the avoidance of error during production) and fluency (real-time rapid language production), with studies of accuracy being less consistent (Sauro & Smith, 2010). In spontaneous synchronous spoken interaction tasks, the affordance lies in offering opportunities to develop fluency, or a learner's capacity to mobilize his or her system to communicate in real time. However, some studies have highlighted learners' perceived deficiencies in modalities resulting in anxiety about speaking with synchronous video or audio tools, which give learners little time to rehearse their statements and which can thus create anxiety (Hampel & Stickler, 2012).

### 3. Purpose of the study

The aim of this study is to explore the relationship between agency and avoidance by focusing on three dimensions of learner choices relating to technological tools, time mode and language use. We aim to understand why some learners may have chosen to avoid the synchronous spoken interaction tasks, which we understand as a form of planning and as a way for learners to exercise their agency. We aim to give some insight for future design for synchronous spoken interaction tasks, and into how online learners may be scaffolded better. The research questions are:

1. How do learners exercise their agency in a synchronous online spoken interaction event in relation to choice and control of a) technological task features; b) language used; c) time modality?
2. How do the choices intersect?
3. What effect do the choices have on time spent in the TL?

### 4. Methodology

#### *Participants*

The participants were students in an English as a Foreign Language (EFL) class as part of their degree programme at a fully virtual university. The learners were in a B2.1 level (lower intermediate) group. Eight adult students, three male and five female, aged between 26–55 years old were engaged in a virtual synchronous peer-to-peer speaking task. Students were bilingual (Catalan and Spanish) with English as an additional language.

#### *Materials*

Learners were asked to complete two previously unseen tasks. They had the same text-based instructions and photo for partners A and B, as well as the means to navigate. Each dyad had to collaborate to complete two out of four tasks on the topic of ‘travelling’. The first was an information gap task (spot-the-difference) and the second was an opinion sharing task. Learners had four differences to find. The second task used the same photo as the first, accompanied by an open question: What kind of activities can people do in a holiday destination like this? General instructions and guidelines were available to students as online documents but some instructions were also offered in text form on the interface screen hosting the task. Tasks were compulsory course assignments and there was no time limitation for tasks.

#### *Design*

The study adopted an exploratory case study approach where students form four dyads. Case studies in language learning facilitate an understanding of learners’ issues, experiences, developmental pathways, insights, or performance within a particular linguistic, educational or social context (Duff, 2014). The dyads came from a data set from a previous study by Appel, Robbins, Moré, and Mullen (2012) that explored the effect of different interfaces within a technological tool in which the results from a learner questionnaire indicated that some learners had scripted or pre-prepared their spoken interaction beforehand. Therefore, **173**



a purposive sampling approach was used to further explore this phenomenon and its effect on spoken interaction. Cases that showed evidence of engaging in spontaneous speech (e.g. interruptions, shorter turns) and those that did not (e.g. longer stretches of speech without interruption or overlap and far fewer turns) were included in the sample.

### *Procedure*

Recordings of peer-to-peer spoken interaction was captured using a plug-in for Skype, a free video and audio conferencing tool, and started from the start of the first task until the end of the second resulting in approximately 23 minutes of data. Recordings were transcribed and converted to a text document. The transcriptions were then coded using the analytical framework of peer group interaction developed by Kumpulainen and Muntanen (1999). This supported a microanalysis of evolving peer interactions, focusing on three analytical dimensions: the functions of verbal interaction, cognitive processing and social processing. The functions of verbal interaction supported analysis of what learners are doing with language using codes such as Informative, Expository, and Organisational. The analysis of cognitive and social processing focused on interactive dynamics as they occurred across the participants. The cognitive processing focused on ways in which students approached and processed learning tasks, highlighting working strategies, situating positions towards knowledge and learning and towards themselves as problem solvers using exploratory or procedural speech, for example. The social processing characterised the social relationship and types of participation in peer groups, such as collaborative and individualistic.

## **5. Results and Discussion**

In order to understand the results and analysis, first we present a descriptive summary of two different general trajectories taken by the four cases outlined in Table 1 below, which highlights how learners used their agency to follow ‘task-as-workplan’ (Breen, 1987) or form their own workplan, trajectory or ‘line of desire’ (Lukin & Du Boulay, 2003).

The descriptions of the two main trajectories were based on the results from learner questionnaires where some learners indicated they had pre-prepared or scripted their interaction. The coding process of the cognitive dimension revealed that the code ‘exploratory talk’ was absent in the transcripts of cases 2, 3 and 4, yet learners would have needed to use exploratory talk to find the differences. The coding process also revealed that in cases 2, 3 and 4, turns for negotiation for organisational purposes were largely absent compared to case 1, who made a number of negotiated turns for organizational purposes. This confirmed that organization had already taken place for cases 2, 3 and 4 before the recordings. Cases 2, 3 and 4 also used the uncommon word “awnings” in their talk, indicating they had looked up the word meaning. In contrast, Case 1 used circumlocution to express their intended meaning when they did not know the lexical item as can be seen in example 2 below.

Other indicators and qualitative differences, noted in the interactions, were also evident in the transcripts, which we will subsequently demonstrate. These indicators were deemed sufficient evidence that covert planning for cases 2, 3 and 4 had taken place and that case 1 had engaged directly in overt synchronous interaction.

Indicators included the complete absence of the L1 in the transcripts of cases 2, 3 and 4, but which is present in the transcript of case 1 as can be seen in the examples below. The

presence of the L1 (as in case 1) is to be expected in the interaction of bilinguals (Macaro, 2006) at an intermediate level, with whom learners share the same L1.

#### Example 1 from Case 1: Spot-the-difference task

M: *¿Pues empezamos en inglés?* [TRANSLATION: Shall we start in English?]

H: OK. I'm ready to start in English

M: OK. Me too.

H: Um er... Who a start? You or me?

M: You can start if you want.

#### Example 2 from Case 1: Spot-the-difference task

M: ...then at the bottom of the building I can see ...mmm... I'm sure it's not called umbrella, but you know that part to make the entrance of stores or something darker?

H: Yes.

#### Example 3 from Case 2: Spot-the-difference task

L: I see on my picture that there are some window shops with some green awnings about these windows...

N: Er, in my picture the awnings are red so it must be the third difference.

Vocalised forms of private speech were also absent, which can also be expected in conversations intermediate level learners as an aid in the mastery of task-related difficulties (McCafferty, 1994). In case 1 private speech occurred when one learner was mumbling the written instructions to himself as shown in example 4 below, before entering into interaction with his partner.

#### Example 4 from Case 1: Spot-the-difference task

M: OK... (mumbles reading the instructions) you can see the same picture. There are four differences. Then I don't know, er, I think this is a tube station, er, because of the signs, er, on the corner of the building...

Expressions of task difficulty were also absent in cases 2, 3 and 4, but present in case 1 as shown in example 5 below.

#### Example 5 from Case 1: Spot-the-difference task

M: Oh yeah yeah yeah yes mine is a little bit white and yours is blue...

H: Yes. Wuff...

M: Wuff...

H: Very tough

Table 1 below outlines the trajectories of the four dyads until the start of task one. Case 1 and their choices (choices 1–3) reveal one trajectory, evident in the transcripts. The covert trajectories regarding choices of cases 2, 3 and 4 are unknown, so they are considered as one trajectory. We recognise, however, that each dyad and learner will approach the task in unique ways. The description of the trajectories of cases 2, 3 and 4 is based on the fact that they had a common need to make certain choices, and acted on them in order to get to the point of recording. It is a trajectory that is not overtly evident in the transcripts, apart from the indicators that planning had occurred and which are described above. Therefore, choices 1–7 of cases 2, 3 and 4 are descriptions of covert choices taken before recording started, but are not necessarily an exact representation of the actual order of choices taken.

Table 1. Dyad trajectories until the start of the first task

Case 1	Cases 2, 3 and 4
1. Choose not to look at the answers/press the button with the technological tool to start recording	1. Choose to look at the answers with the technological tool. The spot-the difference task is made redundant because learners have seen answers. Now the task is how to perform the task as if they had not looked at answers.
2. Choose language to start interacting in (L1)	2. Choose language for discussion in the planning stage (L1 Spanish and/or Catalan and/or TL)
3. Choose when to change language code to TL (Spanish to English)	3. Choose (negotiate) time mode preference to manage the interaction for planning (synchronous/asynchronous or both)
Learners continue to sustain synchronous spoken interaction in English while attempting to resolve information gap task in TL (English)	4. Choose whether to script or prepare performance in L1 and/or TL (exact strategies unknown)
	5. Choose to look up an uncommon word in English “awnings” (one of the differences in the photo)
	6. Choose how to carry out the performance (language choice(s) unknown)
	7. Choose to perform the task completely in the TL

1a) How do learners exercise their agency in relation to choice and control of technological task features?

Case 1 followed ‘task-as-workplan’ (Breen, 1987) which involved pressing the button to invite their partner to start the task, and checking answers during task-processes with virtual buttons in the tool. Navigational choices in relation to the tool highlight learners’ agentic actions using their “motor systems” (Bandura, 1999, p. 4). Their interaction results in a total of 11 minutes 43 seconds in synchronous spoken interaction, with two seconds interacting in the L1. Cases 2, 3 and 4 did not follow ‘task-as-workplan’ but instead followed their own workplan. They chose to use the tool features to look at the answers, resulting

in much less time interacting synchronously in the TL: between 1 minute 38 seconds and 3 minutes 34 seconds. This choice, however, afforded dyads both 'strategic planning time' (i.e. determination of content and code) to be presented and 'rehearsal time' (i.e. a practice run through of the task) (Ellis, 2005). It is not known to what extent and how much time was dedicated to either type of planning time.

In addition, other foci for planning apart from rehearsal and strategic planning (Ellis, 2005) would have needed to be present for cases 2, 3 and 4 in order for the participants to have completed the task. Although they are unknown, the planning foci would need to have included a number of choices: to choose whether to follow 'task-as-workplan' or to look at the answers; to choose how to manage the task once they had looked at the answers (e.g. to script or prepare, and which time mode and tools would be used to communicate with each other); to choose who was going to say what and in what order depending on the degree of scripting vs. spontaneity they had planned and to decide on the co-ordination of their actions around how and when to record their interaction and to choose how to collaborate together so that it appeared that their recorded interaction was spontaneous. Based on this deduction, the number of choices for cases 2, 3 and 4 in relation to technology and management of their language performance appear to be extended and more complex than for case 1. Although their time interacting synchronously in the TL for the pedagogical task is less than in case 1 (see Table 2), their covert interaction and related behaviour means that they spent some considerable time managing the task in their own way. We cannot confirm, however, whether they were more active verbally offline (as in Suzuki's study in 2013) or if they interacted orally/textually, or synchronously/asynchronously during that time.

1b) How do learners exercise their agency in relation to choice and control of language used?

The two trajectories revealed learner language choices as being different, but both trajectories relate to L1 use for organisational purposes or 'task management' as observed by Macaro (1997). Case 1 codeswitches from Spanish to English in order to negotiate an organizational move whilst simultaneously starting the task. The fact that case 1 codeswitched in the recording (knowing the teacher will evaluate it) implies a positive attitude to L1 use. The amount of interaction time in the TL was greater than in the other dyads. This supports Levine's (2011) suggestion that if a plurilingual approach is employed (in this case by learners during the task itself) learners can have more interaction time in the TL, because they are generally talking more. The code switch supported task processes (collaboration and negotiation) and task outcomes (task completion and time interacting in the TL). However, if cases 2, 3 and 4 also employed their L1 in the pre-task planning we could also say that a plurilingual approach was employed by learners leading to a final monolingual output (Dooly, 2011; Hafner, Li & Miller, 2015)

Cases 2, 3 and 4 do not use their L1 during the recordings. Their organisation of the task has occurred in the planning stage. We do not know what language(s) they used in pre-task planning, but the fact that the L1 is totally absent in the recordings can be explained by beliefs or attitudes towards L1 use (Hall & Cook, 2012) or not needing to use the L1 for language purposes such as lexical problem solving. The second explanation, however, is unlikely, given their language level; some L1 use was to be expected because in learning activities, the language of thought for all but the most advanced user of the language is inevitably in their L1 (Macaro, 2006).

It is unknown whether cases 2, 3 and 4 used their multilingual repertoires to control two (or more) languages in authentic communication (Kramersch, 2009) in the planning

stage. While preparing their performance they may have written their preparatory work, individually or together or in a combination of both. This process, however, suggests some involvement of both their L1 and TL. The choice to use different languages for different purposes can be understood as a strategic choice by learners as found by Hafner, Li and Miller (2015) with TL use being used for the final product (Musk, 2014; Hafner, Li & Miller, 2015).

Cases 2, 3 and 4 prepared their contributions to perform in English rather than to completely abandon the task, suggesting that the avoidance strategy allowed them to scaffold their orientation towards task completion while working on the performance. Importantly, then, task avoidance is not a final choice but rather temporary until they start recording and begin the pedagogical task. The time it took for these cases to script or prepare for the recording is unknown. Despite this, a sustained amount of time is implied by the choices they would have needed to negotiate alongside the scripting and planning of the content itself. From this perspective, L1 use in the planning stage can be considered an efficient language choice as found in studies by Musk (2014) and Üstünel and Seedhouse (2005) to get the job done (Musk, 2014) as they manage the task implying the use of language for management, which they have not studied (Macaro, 2006). Case 1 also uses the L1 within a codeswitch also related to task management purposes.

Furthermore, the recordings of the interactions had to be sent to a forum and shared with other peers. Therefore, the choice to avoid a direct recording of their spontaneous interaction being made public (including potential errors and conversation in the L1) may be explained by any one or more of the performance anxieties: a fear of negative evaluation, communication apprehension or test anxiety (Horwitz et. al 1986).

1c) How do learners exercise their agency in relation to choice of time modality?

Case 1 engaged in the synchronous task mode while cases 2, 3 and 4 reconfigured the time mode to asynchronous or, alternatively, a hybrid version. Although the asynchronous mode afforded learners planning time, the choice to reconfigure the time mode meant that they also reconfigured the language outcome from shared rapid real-time spoken interaction, to speech resembling individual spoken production. In doing so, they reconfigure a task designed to develop fluency into one in which they could focus on accuracy and complexity in their utterances (although whether they achieved it or not is not measured). Given that the recordings were to be evaluated by teachers, this reconfiguration is perhaps not surprising, and highlights Ellis' (2005) observation that context, such as test conditions, can shape planning. The asynchronous time mode afforded rehearsal time and strategic planning. It also afforded time for non-linguistic aspects of the task such as social reasons (Ellis, 2005) or technologically related reasons such as tool familiarity. Furthermore, the mode gave learners a way of reducing any potential anxiety because they were no longer "on the spot" (Hurd, 2007, p. 13).

2) How do these choices intersect?

The results suggest that learners' overarching choice to carry out the 'task-as-workplan' or to work according to their own workplan affected other subsequent choices. Choices relating to the three dimensions of agency analysed suggest that each dimension was (re) configured in order to carry out the task according to dyads' own workplan. Although the exact trajectories and choices of case 2, 3 and 4 during the planning stage are unknown, what is known is that the relationship between the choices across the three dimensions was mutually supportive. The choices were made in pairs, confirming that avoidance as a choice

can be collaboratively co-constructed by participants... as a topic of interaction (Markee, 2011) as well as non-avoidance. The mutual use of both the L1/TL and the ability to interact in a non-recorded asynchronous or synchronous time mode afforded learners time to plan the task in their own way until they were willing and ready to engage in synchronicity in the TL. Although avoidance in cases 2, 3 and 4 is not a communication strategy as in Tarone's (1981) typology, it can be considered a communication strategy of a kind: learners controlled the tool and time-mode in order to secure temporary avoidance, which may be an emerging communication strategy specific to SMC spoken interaction tasks.

### 3) What effect do these choices have on time spent in the TL?

We do not know if and how much time was spent interacting orally in the TL in the planning stage. Case 1 had the longest time interacting in the TL and cases 2, 3 and 4 resulted in minimal time interacting in the TL, as seen in Table 2 below.

Table 2. Total interaction time in the target and first language

	Interaction time in L1	Interaction time in TL (spot-the- difference)	Interaction time in TL (open question)	Total Interaction time in TL
<b>Case 1 Dyad (answers not seen)</b>	2 seconds (in task 1 spot-the-difference)	9 mins 13 seconds	2 minutes 28 seconds	11 mins 41 seconds
<b>Case 2 Dyad (answers seen)</b>	unknown	2 mins 18 seconds	2 minutes 49 seconds	5 mins 7 seconds
<b>Case 3 Dyad (answers seen)</b>	unknown	1 minute 38 seconds	49 seconds	2 minutes 27 seconds
<b>Case 4 Dyad (answers seen)</b>	unknown	3 minutes 34 seconds	59 seconds	4 minutes 33 seconds

## 6. Conclusion

The findings highlight that there may be a whole host of factors as to why cases 2, 3 and 4 avoided the 'task-as-workplan', but they remain unknown. The analysis revealed that reasons for task avoidance (or choosing to plan) transcended the cognitive explanations that have been a primary research focus in the field of SLA. Avoidance as a learning strategy (Musk, 2014) should also take into account other possible social and contextual factors. These include a possible need to be familiar with a peer (who is effectively a stranger) to establish a shared perspective on the task and/or to establish group identity. L1 use may have been used to establish group cohesion as found by Hafner, Li and Miller (2015), or to induce a positive affective role as found by Rolin-Ianziti and Varshney (2008), particularly if anxiety was present. The lack of visibility of paralinguistic features (such as gaze and gesture), specific to audio-based synchronous communication, may also heighten the need to engage in peer familiarisation as an attempt to compensate for the lack of body language (Hampel & Hauck, 2004) or a way to counter the effects of a communication process that can be depersonalized (LeCourt, 1999).

Contextual factors may include the online modality. Given that online learners report **179**

that they have difficulties in finding time to invest in activities (Romero & Gentil, 2014), learners may be inclined to approach tasks economically and efficiently (Batstone, 2005) in order to get the job done quickly. Similarly, conceptualising online language learning as a social practice, which is concerned with situating language and learning practices in the broader social practices of everyday life (Burton & Potts, 2013), might highlight that these learners' typical social practice is to collaborate asynchronously in other subjects in their L1, and to prepare and produce a joint outcome in the form of projects and assignments. Therefore, planning can be understood as an extension of their normal social practice. Familiarisation of both the technological tool and the peer with which they will carry out the task may contribute to an increased sense of technical expertise with the tool, preparing themselves for the experience beforehand and being supported through it (Hampel & Hauck, 2004).

In relation to technology, Kress and van Leeuwen (2001) emphasise that multimodal technology makes demands on learners who have to operate several modes (e.g. speech and writing). Indeed, Satar and Özdener (2008) concluded that the (multi-)modality provided by various CMC tools should be considered in relation to divergent groups of learners, especially when differences such as anxiety and proficiency levels are at stake. The task in this study involves learners managing speech, images (icons) and text (instructions) relating to pedagogical and technological aspects of the task alongside navigational demands requiring physical touch. This expanded multimodal scenario may induce more anxiety or need for familiarisation, for example. Alternatively, students may want to use one mode or space for talk between themselves, which is a different mode or space to the one the teacher uses, a behaviour identified by Hampler and Stickler (2012). Testing conditions (Ellis, 2005) may have also been a factor considering the interaction was to be evaluated.

Affective factors may have included fear of negative evaluation, communicative apprehension, test (evaluation) anxiety, anxiety of the synchronous mode (Hampel & Hauck, 2012) or anxiety caused by tool use. Individual differences may also have been a factor including perceived or actual low proficiency level as found by Musk (2014), asynchronous learning styles as identified by Shahabadi and Uplane (2014), and a need for a private space (Suzuki, 2013) and for practice or being 'face sensitive' (Batstone, 2005; Aston, 1986). Considering that learners were going to be evaluated on their interaction, learners may have focused on accuracy over fluency, seeking to avoid making errors as a face-saving strategy. Learner beliefs regarding L1 use during tasks may have also been at play.

Other factors the authors believe may have been present include learners having a different perceived task outcome to the one intended; they may have perceived that it was important to find all the four differences correctly, rather than engage in the interaction in order to do so.

The covert behaviour is central to this study and has consequences for learning (Suzuki, 2013). Results do not wholly support the beneficial role of covert behaviours in other CMC studies however, as found in Sauro and Smith (2010) and Smith (2008). The difference may exist because these studies focused on synchronous text-based CMC, which afforded learners some think-time during interaction. Similarly, factors in Suzuki's (2013) study, although focused on synchronous oral-based CMC, may also have afforded students think-time as the interaction occurred in a teacher-led, multi-party class of learners, possibly giving learners more think-time than peer-to-peer interaction in pairs. Although all of these technological-interactional configurations may aim and help to develop learners' fluency in real-time (or as near to real-time interaction as possible), the presence of think-time may

modify the on-the-spot pressure felt by learners whilst interacting. Although think-time can be seen as an affordance in these studies, because learners could participate and maintain their ability to interact, it may not be desirable if the aim of the task is to simulate real-time “on the spot” (Hurd, 2007, p.13) rapid interaction. Real-time interaction is what many learners face in one-to-one, face-to-face non-online scenarios, given that this is where most spoken interaction occurs. The learners’ choice to avoid rapid-real time interaction in order to potentially gain think-time therefore, may be to their detriment as it eliminates the valuable practice of real-time interaction solely with another person. Suzuki (2013), in her discussion of the pilot study for her main study questions whether regular, active off-screen behaviour in an online course setting provides merit for L2 acquisition. If time spent spontaneously interacting in the TL in the overt stage is a measurement of task success and is significantly less than in the planning stage, then we might conclude that some off-task behaviour (looking at answers and scripting) is not always beneficial. This conclusion is despite the fact that research on pre-planning suggests that pre-planning results in more fluent and more complex language (Foster & Skehan, 1996).

Once the reasons for avoidance are known, ways to scaffold learners could be introduced. Anxiety appears to be the most probable cause, given that the task involved many anxiety-inducing factors. Offering time to rehearse the same task type rather than exactly the same task may support familiarisation with their peers and task type, and tools to help “bridge the gap” (Hurd, 2007, p. 15) between private practice or rehearsal and the real-time task. This can “promote comfort and confidence” (Hurd, 2007, p. 15). Reassurance for learners who are less proficient, less confident or fear negative evaluation for L1 use may be needed. The evaluation of the interaction also need not take place until lots of practice has occurred and the tool use becomes normalised. Signalling a shift towards the importance of task process rather than task completion (finding the answers) may also need to be emphasized in courses. Introducing peer-to-peer evaluation may also help to move the focus away from the task as a performance to the quality of the task process, helping to reduce any performance-related anxiety related to teacher and other peers listening to them.

## 7. Implications

With such a range of opportunities for agency across the three dimensions, learners can almost completely reconfigure the terms and conditions in which they prefer to and are able to interact. Not all learners’ choices relating to technology, language or time-mode lead them to time interacting synchronously and spontaneously in the TL, which is necessary for developing fluency. Despite the fact that these online learners work primarily within an asynchronous mode, it is still important to provide opportunities for fluency because they offer the conditions that resemble (as closely as possible) the real time face-to-face interactions in the TL that many learners face in their non-virtual lives. The task time mode, however, may not match the learning style of many online learners who have chosen or prefer asynchronicity. If this is the case, this creates a particular challenge for designers of synchronous tasks. Indeed, this study brings into question whether one can effectively design fluency activities that are going to be evaluated by the teacher in the context of an online course in EFL. Further research is needed to identify factors contributing to the temporary avoidance of synchronous speaking tasks. Contextual, affective, social or individual factors need consideration as well as cognitive factors; these may be technology or tool-related, and specific to online learners.



## Acknowledgements

A big thankyou to Doris Stanger, Martyn Baker, Diego Blasco, Dean Wang, Suzanne Córdova Martínez and the Institute for Multilingualism, UIC, for support with writing this paper.

## References

- Ahearn, L. M. (2001). Language and Agency. *Annual Review of Anthropology*, 30, 109–137.
- Appel, C., Robbins, J., Moré, J. & Mullen, T. (2012). Task and tool interface design for L3 speaking interaction online. In L. Bradley & S. Thouësny (Eds.), *CALL: Using, Learning, Knowing, EUROCALL Conference, Gothenburg, Sweden, 22–25 August 2012, Proceedings* (15–19). Dublin: Research-publishing.net. DOI: 10.14705/1rpnet.2012.000019.
- Aragon, S. R., Johnson, S. D., & Shaik, N. (2002). The influence of learning style preferences on student success in online versus face-to-face environments. *The American Journal of Distance Education*, 16(4), 227–243.
- Aston, G. (1986). Trouble-shooting in interaction with learners: the more the merrier? *Applied linguistics*, 7(2), 128–143.
- Bandura, A. (1999). *A social cognitive theory of personality*. In L. Pervin, & O. John (ed.), *Handbook of personality* (pp. 154–196). New York, NY: Guilford Publications.
- Barton, D., & Potts, D. (2013). Language learning online as a social practice. *TESOL Quarterly*, 47(4), 815–820.
- Batstone, R. (2005). Planning as discourse activity: A sociocognitive view. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 277–295). Amsterdam: John Benjamins.
- Benbunan-Fich, R., & Hiltz, S. R. (2003). Mediators of the effectiveness of online courses. *IEEE Transactions on Professional Communication*, 46(4), 298–312.
- Blake, R. J. (2005). Bimodal CMC: The glue of language learning at a distance. *CALICO Journal*, 22(3), 497–511.
- Boonsue, W., Janssem, A., & Srinaowaratt, S. (2015). Interactional Patterns in Face-to-face and Synchronous Computer-mediated Communication in Problem-based Learning Contexts. *Journal of Language Teaching and Research*, 6(1), 99–110.
- Borràs, E., Canals, L., Dooly M., Moore, E., & Nussbaum, L. (2009). DYLAN WorkPackage 3: Educational Systems. Deliverable 4: Creativity & Innovation. Unpublished manuscript.
- Breen M. P. (1987). Learner contributions to task design. In C. Candlin, & D. Murphy (Eds.), *Language learning tasks. Lancaster practical papers in English language education*, Volume 7 (pp. 23–46). Englewood Cliffs, NJ: Prentice Hall International.
- de la Colina, A. A., & García Mayo, M. P. (2007). Attention to form across collaborative tasks by low-proficiency learners in an EFL setting. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 91–116). Clevedon: Multilingual Matters.
- Dewaele, J. (2009). Individual differences in second language acquisition. In W. C Ritchie, & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 623–646). Bingley, UK: Emerald Group Publishing.
- Dooley, M. (2011). Divergent perceptions of telecollaborative language learning tasks: Task-as-workplan vs. task-as-process. *Language Learning & Technology*, 15(2), 69–91.

- Duff, P.A. (2014). Case study research on language learning and use. *Annual Review of Applied Linguistics*, 34, 233–255.
- Ellis, R. (Ed.). (2005). *Planning and Task Performance in a Second Language*. Amsterdam: Benjamins.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second language acquisition*, 18(03), 299–323.
- Garcia, O. (2009). Education, multilingualism and translanguaging in the 21st century. In A. K. Mohanty, M. Panday, R. Phillipson, & T. Skutnabb-Kangas (Eds.), *Multilingual Education for Social Justice: Globalizing the local* (pp. 140–158). New Delhi: Orient BlackSwan.
- Gass, S., & Mackey, A. (2006) Input, interaction, and output: An overview. *AILA Review*, 19 (1), 3–17.
- González-Lloret, M. (2011). Conversation analysis of computer-mediated-communication. *CALICO Journal*, 28(2), 308–325.
- Hafner, C.A., Li, D.C., & Miller, L. (2015). Language Choice Among Peers in Project-Based Learning: A Hong Kong Case Study of English Language Learners' Plurilingual Practices in Out-of-Class Computer-Mediated Communication. *Canadian Modern Language Review*, 71(4), 441–470.
- Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language teaching*, 45(03), 271–308.
- Hampel, R., & Hauck, M. (2004). Towards an effective use of audio conferencing in distance language courses. *Language Learning & Technology*, 8(1), 66–82.
- Hampel, R., & Stickler, U., (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(2), 116–137.
- Hurd, S. (2007). Anxiety and non-anxiety in a distance language learning environment: The distance factor as a modifying influence. *System*, 35(4), 487–508.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern language journal*, 70(2), 125–132.
- Jepson, K. (2005). Conversations – and negotiated interaction – in text and voice chat rooms. *Language Learning & Technology*, 9(3), 79–98.
- Kim, H.Y. (2014) Learning opportunities in synchronous computer-mediated communication and face-to-face interaction, *Computer Assisted Language Learning*, 27(1), 26–43.
- Kitade, K. (2008). The Role of Offline Metalanguage Talk in Asynchronous Computer-Mediated Communication. *Language Learning & Technology*, 12(1), 64–84.
- Knight, S. (1996). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal*, 78(3), 285–299.
- Kötter, M. (2003). Negotiation of meaning and codeswitching in online tandems. *Language Learning & Technology*, 7(2), 145–172.
- Kramsch, C.J. (2009). *The multilingual subject: What foreign language learners say about their experience and why it matters*. Oxford: Oxford University Press.
- Kress, G.R., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction*, 9(5), 449–473.

- Leahy, C. (2004). Observations in the computer room: L2 output and learner behavior. *ReCALL*, 16(1), 124-144.
- LeCourt, D. (1999). Writing (without) the body: Gender and power in networked discussion groups. In K. Blair, & P. Takayoshi (Eds.), *Feminist cyberscapes: Mapping gendered academic spaces* (pp.153-175). Stamford, CT: Ablex.
- Levine, G. S. (2003). Student and instructor beliefs and attitudes about target language use, first language use, and anxiety: Report of a questionnaire study. *The Modern Language Journal*, 87(3), 343-364.
- Levine, G. S. (2011). *Code choice in the language classroom*. Bristol: Multilingual Matters.
- Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lukin, R., & Du Boulay, B. (2003). Lines of desire: the challenges of interactive educational TV. In 11th international conference on artificial intelligence in education (AIED2003). Sydney, Australia.
- Macaro, E. (1997). *Target language collaborative learning and autonomy*. Clevedon: Multilingual Matters.
- Macaro, E. (2006). Codeswitching in the L2 classroom: A communication and learning strategy. In E. Llorca (ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession* (pp. 63-84). Amsterdam: Springer.
- Markee, N. (2011). Doing, and justifying doing, avoidance. *Journal of Pragmatics*, 43(2), 602-615.
- Martin, J. (2004). Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, 39(2), 135-145.
- Masats, D., Nussbaum, L., & Unamuno, V. (2007). When activity shapes the repertoire of second language learners. In L. Roberts, A. Gürel, S. Tatar, & L. Marti (Eds.), *EUROSLA yearbook 7* (pp. 121- 147). Amsterdam: John Benjamins.
- McCafferty, S. G. (1994). The use of private speech by adult ESL learners at different levels of proficiency. In J. P. Lantolf, & G. Appel (Eds.), *Vygotskian Approaches to Second Language Research* (pp. 117-134). Norwood, NJ: Ablex.
- Mercer, S. (2011). Understanding learner agency as a complex dynamic system. *System*, 39(4), 427-436.
- Miller, E. (2012, Sep 2). Chapter proposals: Agency and SLA. <http://linguistlist.org/issues/23/23-3666.html>
- Moeller, A. J., & Roberts, A. (2013). Keeping it in the target language. In S. Dhonau (Ed.), *The 2013 Central States Conference on the Teaching of Foreign Languages* (pp. 21-38). Eau Claire, WI: Crown Prints.
- Musk, N. (2014). Avoiding the target language with the help of google: Managing language choices in gathering information for EFL project work. *TESOL Quarterly*, 48(1), 110-135.
- Myers Scotton, C. (1983). The negotiation of identities in conversation: A theory of markedness and code choice. *International journal of the sociology of language* (44), 115-136.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- O'Dowd, R. (2000). Intercultural learning via videoconferencing: A pilot exchange project. *ReCALL*, 12(1), 49-61.

- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second language acquisition*, 21(1), 109–148.
- Rolin-lanziti, J., & Varshney, R. (2008). Students' views regarding the use of the first language: An exploratory study in a tertiary context maximizing target language use. *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 65(2), 249–273.
- Romero, M., & Gentil, C. (2013). Characterizing Online Learners' Time Regulation: Comparative Case Studies of Virtual Campuses in France and Spain. In E. Barberà, & P. Reimann (Eds.), *Assessment and evaluation of time factors in online teaching and learning* (pp. 91–110). Hershey: Information Science Reference.
- Satar, H. M., & Özdener, N. (2008) The Effects of Synchronous CMC on Speaking Proficiency and Anxiety: Text vs. Voice Chat. *The Modern Language Journal*, 92(4), 595–613.
- Sauro, S., & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics* 31(4), 554–577.
- Schwartz, D., & Okita, S. (2009). The productive agency in learning to learning by teaching. Unpublished manuscript. [http://aaalab.stanford.edu/papersProductive\\_Agency\\_in\\_Learning\\_by\\_Teaching.pdf](http://aaalab.stanford.edu/papersProductive_Agency_in_Learning_by_Teaching.pdf)
- Shahabadi, M. M., & Uplane, M. (2014). Learning styles and academic performance of synchronous E-learning students. *Asian Journal of Research in Social Sciences and Humanities*, 4(5), 148–161.
- Smith, B. (2008). Methodological hurdles in capturing CMC Data: The case of the missing self-repair. *Language Learning & Technology*, 12(1), 85–103.
- Stickler, U., Batstone, C., Duensing, A. and Heins, B. (2005) Distance and virtual distance. Preliminary results of a study of interaction patterns in synchronous audio graphic CMC and face-to-face tutorials in beginners' language tutorials. In *Proceedings of the 2004 International Symposium on Applied Linguistics and Language Teaching* (pp. 740–750). Beijing and Shanghai: University of Northern Iowa Press.
- Suzuki, S. (2013). Private turns: A student's off-screen behaviors during synchronous online Japanese instruction. *CALICO Journal*, 30(3), 371–392.
- Swain, M. (2006). Linguaging, agency and collaboration in advanced second language proficiency. In Byrnes, H. (Ed.), *Advanced Language Learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.
- Tarone, E. (1978). Conscious communication strategies in interlanguage: A progress report. In H. D. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL '77: Teaching and Learning English as a Second Language* (pp. 194–203). Washington, DC: TESOL.
- Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly*, 15(3), 285–295.
- Tarone, E., & Swain, M. (1995). A sociolinguistic perspective on second language use in immersion classrooms. *The Modern Language Journal*, 79(2), 166–178.
- Tudini, V. (2010). *Online second language acquisition: Conversation analysis of online chat*. London, UK: Continuum.
- Üstünel, E., & Seedhouse, P. (2005). Why that, in that language, right now? Code-switching and pedagogical focus. *International Journal of Applied Linguistics*, 15(3), 302–325.

- Vitanova, G., Miller, E., Xuesong, G., & Deters, P. (2014). Introduction to theorizing and analyzing agency in second language learning: Interdisciplinary approaches. In P. Deters, X. Gao, E. R. Miller, & G. Vitanova (Eds.), *Theorizing and analyzing agency in second language learning: Interdisciplinary approaches* (pp. 91–110). Bristol, England: Multilingual Matters.
- Wang, Y. (2006). Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL*, 18(1), 122–145.
- Wang, K. H., Wang, T. H., Wang, W. L., & Huang, S. C. (2006). Learning styles and formative assessment strategy: Enhancing student achievement in Web-based learning. *Journal of Computer Assisted Learning*, 22(3), 207–217.
- Yang, Y. T. C., & Chang, L. Y. (2008). No improvement – reflections and suggestions on the use of Skype to enhance college students' oral English proficiency. *British Journal of Educational Technology*, 39(4), 721–725.

### Author biodata

**Elena Barberà** is currently Head of the Doctoral Program in ICT and Education of the Open University of Catalonia in Barcelona (Spain). She is also an adjunct professor for the international doctorate in Nova Southeastern University in Florida (USA). Her research activity is specialised in the area of educational psychology, a field in which she has several publications, conferences and educational courses, relating in particular to knowledge-construction processes and educational interaction in e-learning environments, evaluating educational quality and assessing learning, distance learning using ICT and teaching and learning strategies.

**Janine Knight** is currently carrying out her doctorate at the Open University of Catalonia in Barcelona (Spain) in learner agency and online spoken interaction tasks in the field of SLA. She is also a teacher trainer and English teacher at the Universitat Internacional de Catalunya also in Barcelona.



## Navigational acts and discourse: fostering learner agency in computer-assisted language learning

Janine Knight

Universitat Internacional de Catalunya (UIC), Spain

(email: [janine@uic.es](mailto:janine@uic.es)) and

Universitat Oberta de Catalunya (UOC), Spain

(email: [jknight@uoc.edu](mailto:jknight@uoc.edu))

Elena Barbera

Universitat Oberta de Catalunya (UOC), Spain

(email: [ebarbera@uoc.edu](mailto:ebarbera@uoc.edu))

### ABSTRACT

Fostering learner agency is a primary goal of Integrative Computer Assisted Language Learning, a type of CALL (Warschauer, 1996) that encompasses networked learning and multimedia, including hypermedia. Although navigation has been a focus of attention in some more established CALL scenarios such as Intelligent Computer Assisted Language learning (ICALL) systems in second language acquisition (SLA) research, much less focus has been on more emerging CALL scenarios such as augmented reality, amongst others. Therefore, if learner agency is to be fostered, identifying how it manifests is important.

This study focuses on 'directional agency' (Knight, Barberà and Appel, 2017) which is agency exercised in relation to navigational acts during language learning tasks. Using our own case study data as a starting point (where learners carry out navigation as part of the tasks), the paper draws attention to the somatic acts of navigation while using spoken language as part of the multimodal experiences learners can face. A specific literature review was carried out on other current and emerging CALL scenarios that were considered as cases in order to survey the presence of intentional navigational acts and talk, explore and understand it as a phenomenon in the field and to refine directional agency as a construct.

Results suggest that directional agency is present across other CALL scenarios and task types; directional agency can be shared across learners and technological features; there are multiple forms of navigation including embodied navigation and whole or partial human body navigation that can be occurring in the same task. Navigational acts can accompany learner-learner talk and they can also form part of learner-computer 'talk'. Learners and computers can act as "semiotic initiators and responders" (Coffin and Donohue, 2014), resembling sequential turn-taking of talk so that both learners and digital technologies can be understood as potential actors in the task discourse.

**Key words:** learner agency, Computer-Assisted Language Learning (CALL), task-based language learning, talk, navigational acts

## 1. Introduction

Computer assisted language learning (CALL) has become increasingly diversified and sophisticated, encompassing not only online instruction but also gaming, synthetic environments, digital sensory environments, simulations and augmented reality (González-Lloret, 2017a). These current and emerging CALL scenarios potentially signal a shift away from fixed and static learning experiences that has characterised much of formal language learning and could be said to be a sedentary experience: predominantly involving sitting, listening to the teacher and/or other students and perhaps moving around the class. Original CALL technologies gave face-to-face language learning an additional means for language learning, namely the computer with which the search for and study of applications for it became a research focus (Levy, 1997). This included research on pedagogical tasks such as role-plays and information-gap (information swap) where learner talk is a main objective (e.g. oral interaction) or a secondary one (e.g. collaboration).

However, with many current and new technologies used in CALL, such as mobile phones and augmented reality eyewear, learner navigation is a prominent feature within or outside of a screen and suggests movement of some kind. Despite this, a predominant research focus in second language acquisition studies (SLA) is the oral or written language produced by learners. While this is arguably logical, some researchers suggest that research in the SLA field reflects what Block (2013, p.56) calls “the lingual bias”. While learners’ speech as a semiotic mode has been a main focus, there is a lack of research on the impact of multimodal communication in online classrooms (Hampel and Stickler, 2012) and mode in task-based learning in general (Gilabert, Manchón and Vasylets, 2016). The possible result of this for SLA research may mean that the study of navigation processes in tasks, where learners need to utilise their motor and sensory systems as part of their intentional movements in tasks, can be overlooked.

Navigation can be considered part of the somatic mode: “relating to the body, or affecting the body” (Somatic: Mirriam Webster, 2017) and which can fall within the research on mode in the field of SLA. Because of the expanded use in CALL to include mobile devices as well as eyewear technologies, research on talk-in-interaction while moving has been made more relevant in SLA (Thorne, Hellermann, Jones and Lester, 2015). Understanding more deeply about learners’ navigational acts and its relationship with talk may help task designers and teachers to foster learner agency in future tasks and give insight into task based learning that requires movement for navigation and considers its impact on language learning

For this reason, this study seeks to further explore how learners are navigating with current and emerging technologies in language learning tasks. In particular, the relationship between navigation and talk because interaction that involves collaborative, recursive, turn-taking, is central in SLA theory as it is believed to lead to gains in language acquisition (de la Colina and Mayo, 2007). We are interested in learners’ agency in the task process that pertain to learners’ navigational choices and actions. This is to further understand, develop and refine the construct of learners’ “directional agency” (Knight, Barberà and Appel, 2017) identified as a phenomenon in our own data, which we will subsequently come to. We aim to understand how directional agency may appear in different tasks, with different digital tools and CALL scenarios.

First we revisit our own data from a computer-mediated communication scenario (CMC), a type of CALL where



navigation can form part of tasks and which is the background of this study. We then review recent studies in SLA task-based CALL language learning events that encompass a range of current and future CALL scenarios focusing on talk and navigation which form part of a systematic literature review. We then report on the results of the analysis using a social semiotic multimodal perspective, rather than a solely linguistic one, because it is “a theoretical perspective that brings all socially organised resources that people use to make meaning into one descriptive and analytical domain” (Bezemer, et al., 2012, p.1). By analysing how navigation alongside speech works (or does not work) together in the task process, we attempt to take into account the “lingual bias” (Block, 2013, pp.56) in SLA research. We bring into focus learners’ navigational acts pertaining to the somatic mode, carried out by learners in pedagogical tasks.

## **2. Background**

The background for this study starts with our own data on learner agency in task-based CMC that used an audioconferencing tool to facilitate peer-peer spoken interaction. The learners were adults on a language learning course of English in a 100% virtual university. Using a case study approach, Knight, Barberà and Appel (2017) found that learners could not only exercise different types of agency pertaining to speech, but were also making choices about navigation and acting on those choices with the audioconferencing tool, which in turn, was deemed another type of agency. This was called “directional agency”. This type of agency was identified as being physical in nature but also involving speech. Subsequent studies (Knight and Barberà, 2016; Knight, Dooly and Barberà, 2017a; 2017b) found that depending on the navigational trajectories of learners, learner talk about navigational aspects of the task (as a topic of talk) either became part of the talk between learners during the pedagogical task or was absent from their recorded talk. This outcome depended on whether learners allowed the interface to sequence them while talking in the target language of English or not. From these results, three main things in relation to navigation during these tasks were concluded. The first was that learners could reconfigure the officially designed navigational route. Learners could create their own “lines of desire”, an architectural term used by Luckin and Du Boulay (2003, p.1) applied to technology use, that refer to the paths people make, that are often shortcuts ignoring the given route. The “lines of desire” can also change as learners accommodate new content and changes in the environment. In this sense, learners can agree or disagree with an official route or navigational pathway and create their own path instead. Secondly, learners were not only negotiating the task content orally between themselves through oral turn-turn taking, but that they were also negotiating physically with the technological tool by navigating differently to the sequence intended by the designer (Knight, Dooly and Barberà, 2017a). Thirdly, in addition to talk with their peer, learners were also communicating non-verbally with navigational icons or screen-based resources (e.g. clicking on ‘next task’ and ‘close’ buttons) and that this communication resembled an initiation/response sequence typically attributed to oral speech (Knight, Dooly and Barberà, 2017b). Navigation therefore, could not only form part of the oral discourse as a topic of talk with their peers but navigation screen-based resources such as a pop-up could form part of a non-verbal initiate/response sequences with learners, which was understood as a form of multimodal turn-taking (Knight, Barberà and Dooly, 2017b).

## **3. Theoretical Framework**

### **Learner Agency**

Fostering learner agency is a primary goal of Integrative Computer Assisted Language Learning (Integrative CALL), a type of CALL that encompasses networked learning and multimedia, including hypermedia (Warschauer, 1996). Agency in this study is understood as “the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives” (Martin, 2004, p.135) and which is socio-culturally, contextually and interpersonally mediated (Mercer, 2011). In order to carry out intentional actions, learners can use their motor and/or sensory systems (Bandura, 1999) and/or language (Austin and Urmson, 1962). Pedagogical tasks, such as role-plays, information-gap or quests carried out through technology, typically require learners to receive or produce language (spoken or written) as well as mediate with various visual, textual and/or audio screen-based resources. With respect to navigating during tasks, learners may also be required to use whole body (walking) or partial body movements (e.g. hand, click, pointing a device) to achieve task goals.

As aforementioned, “directional agency” (Knight, Barberà and Appel, 2017, p.276) was identified as pertaining to learners’ choices and acts pertaining to navigation and speech. This type of agency is described as being “physical interaction with technology accompanied by spoken interaction”. It is defined as “choices preceding and accompanying physical moves made by learners to e.g. navigate, check or submit answers either individually or collectively using non device related technological features” and as having characteristics involving “individual (physical) interaction” and “joint negotiation (speech/written)”, (Knight, Barberà and Appel, 2017, p.290).

### **Navigation and Task-based CALL**

Gallistel (1990) defines navigation as “the process of determining and maintaining a course or trajectory from one place to another” (Gallistel, 1990, p.35). A central question in many studies about e-Learning environments is how navigation affects learning (Armitage, Wilson and Sharp, 2004). In CALL research, a main focus on navigation has related to Intelligent Computer Assisted Language Learning (ICALL) systems and hypermedia, (Tsirigia and Virvou, 2001; Heift, 2002 and Amaral and Meurers, 2011). However, more recent studies in CALL highlight other ways of navigating shaped by the scenario, the task and the tools being used. These include interface and net-based navigation in CMC, whole-body navigation with mobile devices, and virtual navigation in virtual environments. These scenarios underscore the potentially shifting relationship learners may have with navigation during tasks across different CALL scenarios.

The studies now presented pertain to navigation and talk identified in four CALL scenarios that arose out of a systematic literature review. The review was added to with results from our own studies using CMC audioconferencing. These reflect four current and emerging task-based scenarios namely audio and video CMC, virtual environments, digital sensory environments, simulations and augmented reality. The criteria for the selection of these specific studies are outlined in the methodology section of the paper.

### **Computer-Mediated Communication (video and audio) tasks, talk and navigation**

In audio and videoconferencing tasks, navigation can accompany learner-learner spoken interaction. Using

videoconferencing, Guth and Helm's (2011) study of three tasks using Skype for telecollaboration showed that in some tasks, learners switched from the skype window to an internet browser or kept other tools open on the screen at the same time. In audioconferencing, Levy and Kennedy (2004) highlighted the affordances of an audioconferencing tool to allow two people to browse through web sites together and see the same on-screen material as they talked. Satar and Özdener (2008), in a study of CMC voice chat mention that learners could press buttons to connect and disconnect (which we consider part of navigational processes). In audioconferencing CMC, Appel, Robbins, Moré and Mullen (2012) focused on how navigational features of a tool facilitated spoken interaction, might guide the conversation and act as a scaffold. Also with the Tandem tool, Knight, Barberà and Appel (2017) found that learners could orally negotiate navigational aspects of tasks through CMC and this negotiation can find its way into the spoken discourse of the task between learners. Knight and Barberà (2016) also found that learners carried out non-verbal communication with navigational screen-based resources directly (e.g. clicking on a pop-up, which invited learners to 'close'). Learners therefore, not only communicated with their peer orally but communicated through clicking somatically with screen-based resources. Knight, Dooly and Barberà (2017b) taking up the notion by Coffin and Donohue (2014, p.28) of learners as "semiotic initiators and responders" found that navigational screen-based resources could invite learners to be non-verbal semiotic responders through the clicking of a (navigational) resource. Navigational screen-based resources were understood as initiating and/or responding to learners. This communication with the navigational screen-based resources resembled the sequentiality of talk so that Initiation/response sequences could therefore be understood as a form of multimodal turn-taking.

### **Virtual environments, talk and navigation**

Studies on task-based virtual environments that mention navigation and talk include studies by Zheng, Young, Wagner and Brewer (2009) who discussed moving and talk within a virtual gaming quest. Learners coupled their chat with working (through their avatars) towards a common goal. The control of the instructor was also often negotiated between collaborators. Jauregi, Canto, De Graaff, Koenraad and Moonen (2011) also highlighted that students communicated synchronously with native speakers in the target language while undertaking action together in a *Second Life*, 3D virtual world.

### **Digital sensory environments, talk and navigation**

Seedhouse (2017) outlines the development of a digital sensory kitchen environment where learners acquire language skills by cooking and interacting with a digital smart kitchen that recognizes learners' movements and uses of utensils. Park and Seedhouse (2017) focused on the kitchen's use for vocabulary learning. The computer gave learners verbal instructions while they were carrying out the task, which the authors likened to satellite navigation while driving a car. The digital system aided by digital satellite signal data, tracked learners' progress in the task and provided timely context specific feedback. Whereas sensors were applied to kitchen objects, the instructions of the task came through a tablet and graphical user interface. Learners could also input information into the system by pressing a number of buttons which meant they could navigate the graphical user interface (GUI) going back, repeating, request help, pause, check where they were in the recipe amongst others.

## **Simulations and Augmented Reality, talk and navigation**

Whereas digital sensory environments can imply whole body navigation (e.g. walking), simulations and augmented reality can include whole body navigation as well as virtual navigation, whereby a virtual screen-based thing or person is moved through navigation. The latter typically involves arms, hands and fingers of learners, which is partial body movement.

A study by Perry (2015) on a quest-based learning and augmented reality via mobiles for university students showed how using Global Positioning System (GPS), a university could transform into a virtual francophone world, where students could interact with characters, items, and media as they improved their language skills and discovered their campus. An augmented reality and interactive storytelling platform was used to create a virtual narrative treasure hunt (quest). At each location, players interacted with virtual characters that directed them or gave them quests with clues or options to further the storyline. These interactions took place either in the form of written text or audio and video recordings to which the student had to respond. Quests also involved challenges ranging from taking pictures of specified objects, to collecting virtual objects, to exploring locations on the map. Groups or individuals could select a variety of options to make their own learning path. Slusareff and Boháčková (2016) took a games-based approach to mobile learning, where game interaction was based on the multimedia content triggered by locations or QR codes; when the player approximated to the area of marked GPS coordinates or read a hidden QR code, the application signalled (vibration and sound) that there was new content. Afterwards, players could shift the story through interactive interviews consisting of image (mostly virtual characters connected with the place), text and multiple choice where a player could answer and thus unfolded a customized version of the story. The branching (that allowed for different instructional sequences) empowered players in partly choosing their path and storyline but all possible branches resulted into a unique ending. The players could navigate themselves as well with compass – the augmented reality tool adding a layer to the smartphone camera that showed direction and distance of other locations.

Thorne, et al. (2015) focused on the ways that small groups used digital technology as they moved through a physical environment in order to accomplish a quest-type task during a mobile augmented reality game play. The game was available on one mobile digital device (an Apple iPhone) that was shared by three players as they negotiated a set of point-to-point route finding tasks. The single device was oriented towards by group members in different ways, via talk-in-interaction, as they accomplished the game activity. The practices for talk-in-interaction, included gaze, postural alignment and deictic expressions, and were used by the participants to maintain their constitution as a group, to accomplish a shared visual focus on the single device, and to explicitly transfer the device from one player to another. A distinctive feature was that only one participant in a group of three was carrying a mobile device, which created a dynamic in which the group, together, would orient toward the device. Each small group needed to coordinate their interpretations of the instructions for the activity and their movement around the university campus. The instructions for the activity were available on the device and the device featured a GPS networked map.

## **4. General Objectives**

The main aim of this paper is to further explore the relationship between navigation and talk in four current and emerging CALL scenarios in language learning tasks to see if the phenomenon of directional agency appears in these. We aim to know if and how talk occurs in these scenarios within a broader notion of discourse to include non-verbal communication with the screen or technologies. We are also interested in whether Coffin and Donohue's (2014) notion of semiotic initiations/responses can be used to characterise this discourse. An additional aim is to understand how meaning may be being negotiated beyond the lingual, as this is an important focus in SLA research. We want a deeper understanding of the phenomenon involving talk and navigation in tasks across CALL scenarios in order to extend, develop and refine the construct of directional agency.

Therefore, our research question is as follows:

RQ1 Is there evidence of directional agency manifesting in other recent or emerging task-based language learning CALL scenarios that involve talk, and if so how?

## **5. Methodology**

The study uses a cross-case analysis rationale in order to see if there is transferability of the phenomenon found in our cases across other studies. These studies are considered cases after a process of selection from a systematic literature review which we subsequently describe.

First we undertook a conceptual review of directional agency in order to have a basis as a specific framework to work from, based on navigational acts and talk identified in our own studies. We then carried out a systematic literature review which is the most appropriate research method to provide a reliable overall picture of studies to comment on, extend, or develop theory (Baumeister and Leary, 1997) and to formulate an overarching conceptualization (make a point, rather than summarizing all the points everyone else has made (Sternberg, 1991). However, the two major problems with this method applied to this study were that we were focused on a phenomenon that had only been recently identified and given a name as well as the fact that we were working within a research field with "a lingual bias" (Block, 2013, p.56). For these reasons, studies were not focused on or reporting the phenomenon we were interested in. Another challenge was that the term pedagogical "task" has a very specific meaning and criteria in SLA but at the same time it is a word that is commonly used as a synonym for "activity" or "project". Therefore, the specific literature review needed to be undertaken by someone who was an expert in the SLA field and more specifically, task-based learning. As a base for the systematic review, we used a pre-existing literature review carried out by such an expert in the field of task-based language learning and technology, namely González-Lloret (2017a). This review of current and emerging CALL scenarios was presented at an International conference on Task-Based Language and Teaching in 2017. These included 1) Computer Mediated Communication (CMC), 2) digital sensory environments, 3) simulations and augmented reality and 4) virtual environments, 5) online gaming, 6) online instruction and 7) intelligent computer-assisted language learning (ICALL) systems (appendix A). These last three were omitted from our systematic literature review. Based on personal correspondence with González-Lloret, we established that the literature selection was based on task-based learning and technology and the studies selected by González-Lloret were ones that actually included pedagogical task rather than class activities (even if they

said they were or even if they did not call them so) (González-Lloret, 2017b). From this selection, we then took a progressive focusing approach, which was refined by Stake (1981). This approach involves observation of the site, further enquiry, beginning to focus on the relevant issues and then seeking to explain. We used the model developed by Sinkovics and Alfoldi (2012), focusing on four steps: 1) choosing a sample and context (concerned with moving the research out of theory into the field); 2) data collection and preparation; 3) data analysis and comparison with theory and 4) discussion and write up.

The sample we chose was based on four CALL scenarios from González-Lloret's original review. We included two scenarios which she deemed future CALL scenarios because we wanted to include tasks that were using the newest technologies (digital sensory environments and augmented reality). We also included our own CALL scenario (CMC) and then one more (virtual environments). The context was formal language learning scenarios that used a task-based learning approach mediated with some form of digital technology. The data collection and preparation consisted of first looking for studies that mentioned navigation but where navigation was not the main focus. We included studies after 2009 to the present date but included studies before then if the search results added some value to the current study. We then focused on learner talk. We initially focused on tasks for spoken interaction but given that many tasks involve a number of language skills e.g. speaking and writing, we left that aspect open, and therefore included all studies that mentioned learner talk. However, we soon realised in the collection stage that learners were not the only ones talking. Although directional agency was originally conceptualised as encompassing learner talk in the original construct, we expanded the inclusion of studies that made reference to include just talk, reasoning that learners may be responding to computer-generated talk. This was also based on our own ongoing work that found that learners responded to computer-generated textual requests in tasks, which could be considered 'initiations' (Knight, Dooly, Barberà, 2017b). We included different task types because we were interested in the transferability of the phenomenon. Based on this inclusion/exclusion criteria i.e. talk and navigation, each study was read and selected by the main researcher. Books were also included so the chapters and indexes were scanned for mention of navigation and talk. We added other studies to the selection namely those that had led up to and formed part of our own studies on talk, navigation and agency including Appel, Robbins, Moré and Mullen's study (2012) on navigation and interface versions. Data analysis involved creating a table as a data display (Yin, 2009) for analysing the studies. This was used to support an analysis as to how the studies in the field compared with aspects of the original definition. We also included whom/what the talk was emerging from which we link to initiation/response sequences that are implied (i.e. learners-learners or computer-learner interaction). In addition, we read the studies in order to understand how meaning negotiation may be being made verbally and non-verbally in tasks. The results of this process helped to shape the original construct of directional agency, leading to a revised version, which is presented at the end of this paper.

## **6. Results and Discussion**

The research question pertains to whether there is evidence of directional agency manifesting in other recent or emerging task-based language learning CALL scenarios that involve talk, and if so how.

Before presenting the results, we review the original definition of Directional Agency as defined by Knight, Barberà and Appel (2017, p.290). Directional agency is "learners' choices and acts pertaining to navigation

and speech". It was "physical interaction with technology accompanied by spoken interaction". It was defined as "choices preceding and accompanying physical moves made by learners to e.g. navigate, check or submit answers either individually or collectively using non device related technological features". The characteristics involved "individual (physical) interaction" and "joint negotiation (speech/written)".

In our analysis we compare aspects from the original definition with the selected studies identified from the four CALL scenarios namely 1) its nature; 2) the physical interaction with technology accompanied by spoken interaction; 3) choices preceding and accompanying physical moves made by learners in the navigational process and 4) that it could be carried out individually or collectively using non device related technology.

Results suggest that directional agency appears to be present across different CALL scenarios but the construct needs to be refined in the light of its application to other CALL scenarios.

### **1) The nature of directional agency**

The physical moves to navigate in CMC scenarios typically takes place with human hands and fingers. This is evident in studies on CMC with audio whereby learners click on a screen-based resource/button to look at answers to tasks (Appel, et al., 2012 and Knight, Barberà and Appel, 2017). Learners click on pop-ups that require closing or click on buttons to move learners on to the 'next task' (Knight and Barberà, 2016) or interface page. Learners browse through websites together (Levy and Kennedy, 2004) implying website and webpage navigation. Learners switch from windows to an internet browser or keep other tools open on the screen at the same time in videoconferencing (Guth and Helm, 2011). Learners click to connect and disconnect in voice chat (Satar and Özdener, 2008).

The physical nature in virtual environments is evident in learners working through (moving) their avatars toward a common goal as in the study by Zheng, et al. (2009) and undertaking virtual action together in the form of avatar-embodied interaction (Jauregi, et al., 2011). Moving avatars such as running or flying involves clicking the keys on a keyboard so while the nature is physical it is at the same time embodied (pertaining to the human body) in nature.

In digital sensory environments, learners can move around a physical learning space (kitchen) as well as press navigational buttons to input information into the system (Seedhouse, 2017) therefore highlighting the physical, whole-human body being part of navigation. This also occurs in simulations and augmented reality scenarios, whereby physical moves to navigate can also take place with the whole body as suggested in studies by Perry (2015), Slusareff and Boháčková (2016) and Thorne, et al. (2015). Learners need to walk from one place to another, stopping at different points to then move on to another once again.

The nature of learners' agency across these scenarios is carried out through the relationship between the learner/s body and the tool/s used. The physical nature encompasses human hand-based navigation based on the assumption that hypertext, browser navigation and moving avatars typically occurs in this way. It also encompasses movement of the human whole body in relation with the tool as the tool navigates the body (e.g. through being held or being pointed) in a trajectory as well as the embodiment of humans represented virtually

and moving virtually as avatars. These are not necessarily carried out in isolation from each other but may run parallel with each other in a task, such as moving an avatar (through human hand-based navigation) to move an embodied 'you' or walking in a physical space while also navigating with a screen or other system.

While the nature of directional agency may be experienced physically by learners, learners' intentional navigational moves appear intrinsically inter-connected to tool use. Therefore, directional agency can be considered both physical and digital in nature.

## **2) Physical interaction with technology accompanied by spoken interaction**

Many studies noted spoken interaction between learners, alongside physical interaction with technologies for navigation. In a digital sensory environment, learners move themselves around the kitchen (Seedhouse, 2017) whilst responding to verbal instructions from the computer.

In a simulations and augmented reality scenario, Thorne, et al. (2015) found that the talk between learners was used to share a focus on the device and to explicitly transfer the device between players as well as their oral negotiation of a route from point-to-point. In a virtual environment Zheng, et al. (2009) found that coupled with learners having to negotiate the moving of the avatars, they also had to orally negotiate the controlling role of the instructor between collaborators. What this result points to is that learner-learner talk or computer-learner talk for navigational purposes may be multi-layered, either centring on the navigation of the device and/or resources on the screen and/or centring on the navigation of the person. In addition, the spoken interaction for navigation also may accompany physical interaction between learners, such as the transferring of a device. What these studies have in common is the oral negotiation present between learners about what to do with different objects (e.g. devices, avatars, navigational buttons, themselves) in the navigation process; the negotiation of the route in the process. These results concur with Zheng, et al. (2009) proposal that interaction and actions and movement are negotiated towards learners' goals.

Furthermore, other studies mention how the 'talk' can be carried out by a computer, both verbally to learners as identified in studies by Seedhouse, (2017), Jauregi, et al. (2011) and textually as in studies by Jauregi, et al. (2011) and authors (in press). In addition, computers can talk to learners through movement (such as vibration and sound) as identified in the study by Perry (2015) or through a pop-up as found by authors (in press). This physical interaction with technology can be understood as initiation and response sequences: a form of discourse involving a mixture of verbal and/or non-verbal (e.g. textual, vibration and sound) communication.

We identified how the computer can initiate a response by one or more learners in a number of studies. In a digital sensory environment (Seedhouse, 2017) the computer's clues and options were given through virtual characters connected with the place. In an augmented reality game (Slusareff and Boháckova, 2016), learners could respond with multiple choice. In an augmented quest game with mobiles, virtual characters could direct learners and give verbal instructions that were text, audio or video bases so that learners could physically carry out the task (Perry, 2015). In a CMC task with audioconferencing, pop-ups and navigational screen-based resources initiated learners' response to close a pop-up through physically clicking on it (Knight and Barberà,



2016 and Knight, Barberà and Appel, 2017).

In the light of these findings, it appears that while the original premise for directional agency was based on the fact that spoken aspects can accompany or precede navigation and is the case in many studies, the definition should change in three main ways. Firstly, with respect to who is doing the 'talk': that the spoken interaction is occurring between computer-learner not just learner-learner. Secondly, that other forms of communication should be included in the interaction (e.g. vibration and sound from QR codes, oral instructions from a virtual character). As we can see from many of the studies, the computer can talk/communicate so that the learner/s move/s themselves. It is not just that the device or screen-based button can be moved by the learner in some way (e.g. pressing, clicking).

### **3) Choices preceding and accompanying physical moves in the navigation process**

After comparing the studies, we identify that learners' choices and physical moves do not only pertain to navigation of screen-based resources (e.g. virtual buttons or hyperlinks) alongside partial body navigation (hand/finger) as initially conceptualised. In addition, learners' can make choices and act on these with their whole physical body and/or embodied versions of themselves. In addition, learners pointing devices in different directions as an initiation requiring a technological response can also be considered a navigational act. This results leads us not only to modify the definition in this aspect but rather to see how its application can be expanded.

### **4) Being carried out individually or collectively using non-device related technological features**

While studies attest to individuals, pairs and groups in carrying out directional agency through technologies, what it did not encompass was that in many studies the technologies are not only digital tools for navigational purposes but they could also act as protagonists or digital "others" (Raudaskoski, 1999) through the actions of giving instructions/clues etc. in the task process. In addition, we found that directional agency can be carried out through both non-device related features (e.g. screen-based buttons) as originally proposed, as well as devices or extensions of a device (e.g. mobile phone and QR codes or keyboards).

As a consequence of the results, the original construct of directional agency has been refined through its application with other CALL scenarios and our understanding of the phenomenon has deepened. We now propose that directional agency is both physical and digital in nature, but also has verbal and/or non-verbal communication accompanying or preceding it, necessary for navigational purposes to be carried out. Directional agency involves intentionality, which may be apparent in the choices and physical moves of learners to navigate or check progress in the navigational trajectory. The intentionality of digital tools is also evident in the messages given to learners (e.g. through instructions from virtual characters or pop-ups instructing learners to close them) as well as where the technologies take learners to (a specific virtual page or a specific physical place). It can be carried out individually or collectively by humans and may also include other digital actors. Navigational choices may be able to be negotiated by learners using a range of verbal and non-verbal modes. Navigational acts can be understood as momentary events such as a click on a hyperlink or a larger trajectory of navigating from one place to another with several steps along the way. The physical or

physical/virtual moves that form part of it can involve human, embodied partial body (such as pressing a key on a keyboard) or whole body human or embodied movement (such as walking in virtual or non-virtual environments) along a trajectory.

## **7. Conclusions**

The study has sought to identify, explore and understand the phenomenon of directional agency. This was in addition to further exploring the relationship between navigation and talk and understanding the relationship through a multimodal lens. Directional agency was found to be present across other task types and CALL scenarios but as a construct it needed to be refined. We found that directional agency can be shared across learners and digital technologies and that multiple forms of navigation can occur in the same task. Results suggest that agency is not individual but part of a system of action and it implies negotiation which is not only linguistic but multimodal.

The Coffin and Donohue (2014, p.28) notion of “semiotic initiators and responders” has been key in understanding directional agency as a shared system between learners and technological aspects of tasks. While we found that the initiation/response sequence was useful in characterising what we have come to understand as human-human and human-computer multimodal turn-taking in task discourse involving navigation, the sequence needs further research in order to study the sequences in greater depth.

Finally, the initiations and responses observed in the studies are not necessarily mechanistic but rather, appear to involve some form of negotiation including oral negotiation between learners, physical negotiation between computer and learner and/or embodied virtual negotiation between avatars. Given that meaning negotiation is an important research area in SLA theory and research, further research is needed in order to understand how meaning negotiation in CALL scenarios is carried out multimodally, not just linguistically. These multimodal negotiations may also give further insight into how learner agency or types of agency can be carried out (or not) by learners in order to foster agency in future CALL and non- CALL scenarios for optimal language learning.

## References

- Amaral, L. A., and Meurers, D., 2011. On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. *ReCALL*, 23(1), p.4-24.
- Appel, C., Robbins, J., Moré, J. and Mullen, T., 2012. Task and tool interface design for L3 speaking interaction online. In: EUROCALL Conference,, *CALL: Using, Learning, Knowing, Gothenburg, Sweden, 22-25 August 2012*,. Dublin: Research-publishing.net. DOI: 10.14705/rpnet.2012.000019.
- Armitage, U., Wilson, S., and Sharp, H., 2004. Navigation and Ownership for Learning in Electronic Texts: An Experimental Study. *Electronic Journal of E-Learning*, 2(1), p.19-30.
- Bandura, A., 1999. A Social cognitive theory of personality. In: L. Pervin and O. John, eds. *Handbook of personality: Theory and research*. New York, NY: Guilford Publications. pp.154-196.
- Austin, J. L., 1962. *How to Do Things With Words*. Oxford: Oxford University Press.
- Baumeister, R.F. and Leary, M.R., 1997. Writing narrative literature reviews. *Review of General Psychology*, 1(3), p.311.
- Bezemer, J., Jewitt, C., Diamantopoulou, S., Kress, G. and Mavers, D., 2012. *Using a social semiotic approach to multimodality: Researching learning in schools, museums and hospitals*. [online] Available at: [http://eprints.ncrm.ac.uk/2258/4/NCRM\\_working\\_paper\\_0112.pdf](http://eprints.ncrm.ac.uk/2258/4/NCRM_working_paper_0112.pdf) [Accessed 22 January 2018].
- Block, D., 2013. Moving beyond Lingualism: Multilingual Embodiment and Multimodality in SLA. In: S. May, ed. *The multilingual turn: Implications for SLA, TESOL, and bilingual education*. New York, NY: Routledge. pp.54-77.
- Coffin, C., and Donohue, J., 2014. A language as social semiotic-based approach to teaching and learning in Higher Education. *Language Learning*, 64(1), p.1-212.
- de la Colina A. A. and García Mayo M. P., 2007. Attention to form across collaborative tasks by low-proficiency learners in an EFL setting. In: M. P. García Mayo, ed. *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, pp.91–116.
- Gallistel, C. R., 1990. *The organization of learning*. Cambridge, Mass.: MIT Press.
- Gilabert, R., Manchón, R., and Vasylets, O., 2016. Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, p.117-135.
- González-Lloret, M. (2017a, April). *Tasks in Technology-mediated Contexts*. Keynote presentation at The Seventh International Conference on Task-Based Language Teaching, Barcelona, Spain. 20 April, 2017.
- González-Lloret, M., [marta@hawaii.edu](mailto:marta@hawaii.edu), 2017b. *TBLT conference plenary talk*. [email] Message to J. Knight ([janine@uic.es](mailto:janine@uic.es)). Sent Friday 2 June 2017: 03:49.
- Guth, S., and Helm, F., 2011. Developing multi-literacies in ELT through telecollaboration. *ELT Journal*, 66(1), p.42-51.
- Hampel, R., and Stickler, U., 2012. The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, 24(2), p.116-137.
- Heift, T., 2002. Learner control and error correction in ICALL: Browsers, peekers, and adamant. *CALICO Journal*, 19(2), p.295–313.

- Jauregi, K., Canto, S., De Graaff, R., Koenraad, T. and Moonen, M., 2011. Verbal interaction in Second Life: towards a pedagogic framework for task design. *Computer Assisted Language Learning*, 24(1), p.77-101.
- Knight, J. and Barberà, E., 2016. The Negotiation of Shared and Personal Meaning Making in Spoken Interaction Tasks. In: Pixel (ed.), *Conference proceedings of the 9th International Conference of ICT for language learning* November, 2016. Florence, Italy. libreriauniversitaria. it Edizioni press. Italy. p. 248-252.
- Knight, J., Barberà, E. and Appel, C., 2017. A framework for learner agency in online spoken interaction tasks. *ReCALL*, 29(3), p.276-293.
- Knight, J., Dooly, M. and Barberà, E. (2017a). Multimodal meaning making: navigational acts in an online speaking task. [Unpublished manuscript].
- Knight, J., Dooly, M. and Barberà, E. (2017b). Navigating a multimodal ensemble: mediating turns verbally and non-verbally in online interaction. [Unpublished manuscript].
- Levy, M., 1997. *Computer-assisted language learning: Context and conceptualization*. Oxford: Oxford University Press.
- Levy, M. and Kennedy, C., 2004. A task-cycling pedagogy using audio-conferencing and stimulated reflection for foreign language learning. *Language Learning and Technology*, 8(2), p.50-68.
- Luckin, R. and Du Boulay, B., 2003. Lines of desire: the challenges of interactive educational TV. In: *11th international conference on artificial intelligence in education (AIED2003)*. Sydney, Australia. IOS Press.
- Martin, J., 2004. Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, 39(2), p.135-145.
- Mercer, S., 2011. Understanding learner agency as a complex dynamic system. *System*, 39(4), p.427-436.
- Park, J. and Seedhouse, P., 2017. Sight and Touch in Vocabulary Learning: The Korean Digital Kitchen. In: P. Seedhouse, ed. 2017. *Task-Based Language Learning in a Real-World Digital Environment: The European Digital Kitchen*. London: Bloomsbury Publishing. Ch.10.
- Perry, B., 2015. Gamifying French Language Learning: a case study examining a quest-based, augmented reality mobile learning-tool. *Procedia-Social and Behavioral Sciences*, 174, p.2308-2315.
- Satar, H., and Özdener, N., 2008. The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, 92(4), p.595-613.
- Seedhouse, P., ed. 2017. Introduction. In: P. Seedhouse, ed. 2017. *Task-Based Language Learning in a Real-World Digital Environment: The European Digital Kitchen*. London: Bloomsbury Publishing. pp.3-18.
- Sinkovics, R. R., and Alfoldi, E. A., 2012. Facilitating the Interaction between Theory and Data. Qualitative Research using CADQAS. In: G. Symon and C. Cassell, eds. 2012. *Qualitative Organizational Research: Core Methods and Current Challenges*. London: Sage Publications Ltd. pp. 109-131.
- Slussareff, M., and Boháčková, P., 2016. Students as Game Designers vs 'Just' Players: Comparison of Two Different Approaches to Location-Based Games Implementation into School Curricula. *Digital Education Review*, (29), p.284-297.
- Somatic, 2017. *Merriam-Webster.com*. [online] Available at: <https://www.merriam-webster.com/dictionary/somatic> [Accessed 15 May 2017 ]
- Stake, R.E., 1981. The art of progressive focusing. Paper presented at 65th annual meeting of American Educational Research Association. Los Angeles: CA. April 13-17. pp.1-8.

Sternberg, R. J., 1991. Editorial. *Psychological Bulletin*, 109, p.3-4.

Thorne, S. L., Hellermann, J., Jones, A., and Lester, D., 2015. Interactional practices and artifact orientation in mobile augmented reality game play. *PsychNology Journal*, 13(2–3), p.259–286.

Tsiriga, V., and Virvou, M., 2003. Modelling the student to individualise tutoring in a web-based ICALL. *International Journal of Continuing Engineering Education and Life Long Learning*, 13(3-4), p.350-365.

Warschauer, M., 1996. Computer Assisted Language Learning: an Introduction. In: Fotos S., ed. 1998. *Multimedia language teaching*. Tokyo: Logos International. pp.3-20.

Yin, R. K., 2009. *Case Study Research: Design and Methods*, 4<sup>th</sup> ed. Applied Social Research Series, Volume 5. Thousand Oaks, CA: Sage Publications.

Zheng, D., Young, M. F., Wagner, M., and Brewer, R. A., 2009. Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal*, 93(4), p.489-511.

**Appendix A: Literature review of four CALL scenarios presented by González-Lloret (2017) that formed the basis for the systematic literature review**

1) Audio and video Computer Mediated Communication (CMC) and telecollaboration	Bueno-Alastuey, 2013; Granena, 2006; Loewen and Isbell, 2017; Loewen and Wolff, 2016; Satar and Özdener, 2008 and Yanguas, 2010. Levy and Kennedy, 2004. Müller-Hartmaan (2000); Furstenberg 2001;
--	---

	Helm, 2015; Guth and Helm, 2011; van der Kroon, Jauregi, and Jan, 2015; O'Dowd, 2016 and Van der Zwaard and Bennick, 2014; Sadler and Dooly, 2016; Guth and Helm, 2010; Lewis and O'Dowd, 2016: O'Dowd, 2016;
2) Digital sensory environments	Preston, Balaam, Seedhouse, Kurhila, Kotilainen, Rafiev and Olivier, 2015; Seedhouse, 2017; Seedhouse and Knight, 2016
3) Simulations and augmented reality	González-Lloret, 2003; Perry, 2015; Reeder, 2010; Santos, Taketomi, Yamamoto, Rodrigo, Sandor and Kato, 2016; Slusareff and Boháčková, 2016; Salak and Cakir, 2015 and Thorne, Hellermann, Jones and Lester, 2015.
4) Virtual environments	Canto, de Graff and Jauregi, 2014; Ganem-Gutierrez, 2014; Hampel, 2010; Sadler, 2009, 2012; Sadler and Dooly, 2016; Zheng, Young, Wagner and Brewer 2009 and Jauregi, Canto, de Graff, Koenraad, and Moonen, 2011; Milton, Jonsen, Hirst and Lindenburn, (2012).



## Appendix 2. Codes used pertaining to language functions across three task types

Ten of the labels used in this analysis system were used originally by Kumpulainen and Wray's (2002), which in turn were used by others, e.g. Barnes and Todd, 1977; Halliday, 1975; Phillips, 1985; Tough, 1977). However, the system described above emerged from the data collected with each function corresponding to specific research findings across the three task types. The functions and their corresponding codes are presented.

Language functions	Code	Brief description
Interrogative	Q	Asking questions
Responsive	A	Answering questions
Affective	AF	Expression of personal feelings
	Aff-t	<i>Tiredness</i>
	Aff-d	<i>Desire in relation to topic</i>
	Aff-Rs	<i>Reassurance to partner</i>
	Aff-a	<i>Apology to partner</i>
	Aff-h	<i>Humour</i>
	Aff-s	<i>Surprise</i>
Informative	I	Providing information
	PI	<i>Providing personal information</i>
	IPref	<i>Information on preferences, likes and/or dislikes in relation to topic</i>
	IT	<i>Information related to technological aspects of the task</i>
Judgemental	Ja	Expressing agreement
	Jd	disagreement
	Ju	uncertainty
Organisational	OR	Organising behaviour
Reproductional	RP	Reproducing spoken language
Experiential	E	Expressing personal experiences
Argumentational	AR	Reasoning in oral language
Imaginative	IM	Introducing or expressing imaginative situations
Summative	SUM	Summarising pair activity
Confirmation	CON	Confirmation of message received
Opinion	OP	Opinion without reasoning
Clarification	CI	Clarification request
Greeting and leave taking	G	Greeting or leave taking
	LT	
	PInt	Personal Introductions
Miscellaneous	M	e.g. mumbling unaccompanied by other functions



### Appendix 3. glossary of key terms used in this study

<b>Multimodality</b>	Multimodality is an inter-disciplinary approach that understands communication and representation to be more than about language.
<b>Meaning making (sociocultural theory)</b>	Collaborative dialogue (Swain, 2000, p. 112) through the mode of speech. To negotiate meaning, learners “negotiate the message of the input” (Swain, 2000, p. 98)
<b>Meaning making (social semiotics)</b>	Meaning making occurs in a variety of ways, not only with language. Different semiotic resources, each offer distinct potentialities and limitations. (Jewitt, Bezemer and O’Halloran, 2016)
<b>Negotiation for Meaning (NfM)*</b>  *some researchers refer to the concept as Negotiation of meaning (NoM)	Negotiation for Meaning focuses on learners’ oral turns for clarification requests and comprehension checks, among others, in order to overcome comprehension difficulties (Varonis and Gass, 1985) The total number of <i>negotiated turns</i> in a <i>negotiation of meaning</i> sequence are measured, in order to ascertain how various factors e.g. task types can stimulate greater amount of talk in the target language.
<b>Negotiation for meaning (Communities of practice theory)</b>	How we experience the world and our engagement in social processes that are meaningful. Participation is an element in the negotiation process that requires active involvement by learners.
<b>Mode</b>	A set of socially and culturally shaped resources for making meaning (Kress, 2014). Examples of modes include writing and image on the page, extending to moving image and sound on the screen, and speech, gesture, gaze and posture in embodied interaction.
<b>Navigational act (human intention)</b>  <b>(computer generated)</b>	The intentional clicking of screen-based resource by humans for navigation, that when combined with other clicks, forms a trajectory of use. The notion can be expanded to include intentional pointing of a digital resource for navigation (see paper 7).  The term can also pertain to how computers intentionally seek to ‘move’ learners by, for example, presenting humans with a pop-up that humans must ‘click’ to close and which moves humans to another page/step in the navigational path.
<b>Active agent</b>	An active agent takes action on its own when it sees an opportunity for doing so. Examples of active agents are pop-ups. (Fischer and Morch, 1988) for domain-oriented design environments. In this study, the term pertains to the screen-based resources on the interface.
<b>Passive agent</b>	A passive agent waits until it receives a request to act. Both kinds of agents follow a set of rules for when to act (precondition) and what to say (intervention strategy). Fischer and colleagues (1985) were the first to introduce the distinction between passive and active agents (later referred

	to as 'critics'). In this study, the term pertains to the screen-based resources on the interface.
<b>Higher Level Action</b>	A term used to refer to large scale actions, such as a meeting, and is made up of a multiplicity of chained lower-level actions (Norris, 2004)
<b>Lower Level Action</b>	A term used to refer to smaller-scale actions, for example, gestures or gaze shifts that become chains of lower level actions (Norris, 2004)
<b>Semiotic budgets</b>	A term coined by Van Lier (2000, p.252) that pertains to the language classroom, e.g. materials, that provide learners with affordances that have been designed to stimulate engagement with the new language.
<b>Screen-based resources</b>	Resources that pertain to the task interface that are intended for learning purposes e.g. on-screen clues such as photos/videos, various textual clues lead the participants to perform web search activities (Balaman and Sert, 2017)a. . Screen-based resources can be considered part of a learners 'semiotic budget' that specifically pertain to the screen.
<b>Digital literacy</b>	While there is a range of definitions for digital literacy, it is used to refer to the use of digital tools and resources to foster the autonomous development of the basic linguistic and intercultural skills required to engage in interactions. (ECML, 2018)
<b>Learner agency</b>	In this study we use Martin's (2004) definition: as "the capability of individual human beings to make choices and act on these choices in a way that makes a difference in their lives" (Martin, 2004, p.135), while also acknowledging that agency is "socioculturally mediated" (Ahearn, 2001, p.112).